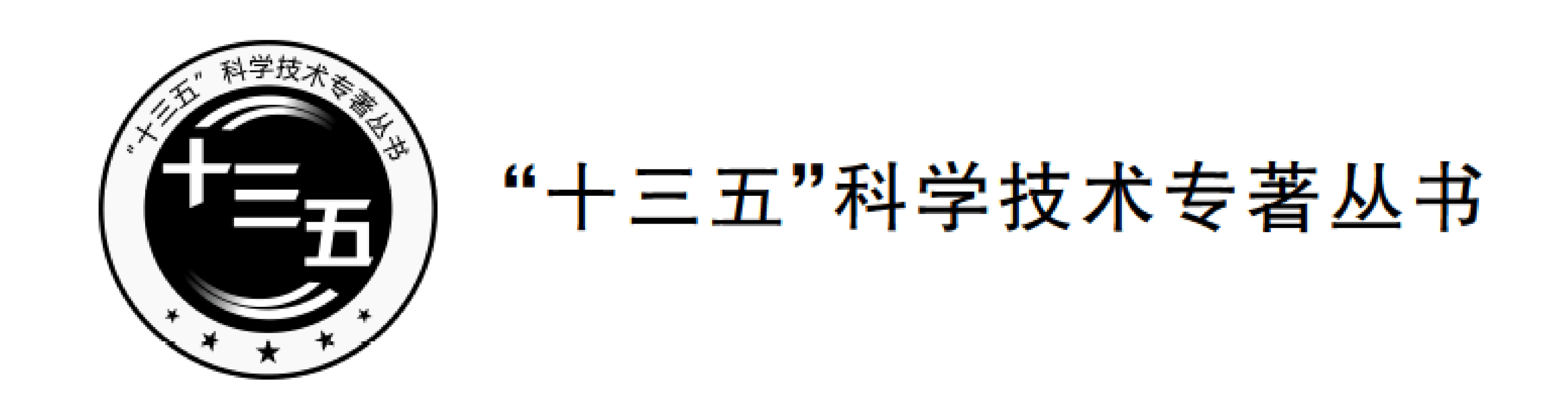


视觉跟踪中的马氏链蒙特卡洛方法

周修庄　鲁继文 著

(MCMC Sampling for Visual Tracking)

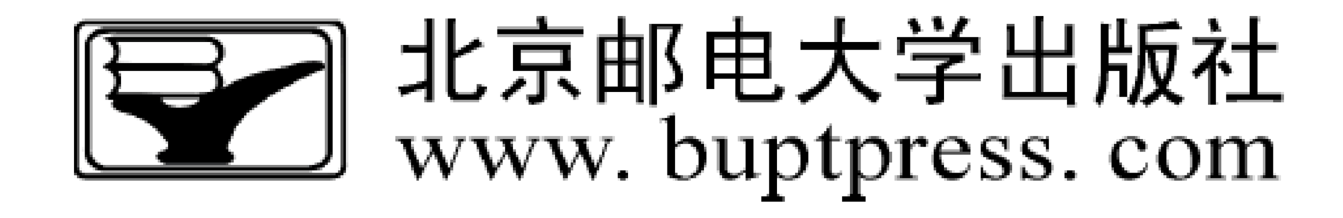
北京邮电大学出版社
www.buptpress.com



视觉跟踪中的马氏链蒙特卡洛方法

（MCMC Sampling for Visual Tracking）

周修庄　鲁继文　著



内 容 简 介

本书基于统计机器学习的理论工具——马氏链蒙特卡洛(MCMC)采样理论,系统地阐述了目标运动不确定性条件下的视觉目标跟踪问题、方法和技术,并结合多种实际应用场景给出算法实现框架和实验分析。全书共4章,第1章介绍了视觉目标跟踪的意义和基本方法。第2章介绍了MCMC采样、自适应MCMC采样的基本理论、方法、实现技术及其在计算机视觉中的应用。第3章介绍了视觉目标跟踪中的序贯粒子采样方法和序贯随机逼近蒙特卡洛采样方法。第4章介绍了面向运动突变视觉目标跟踪的自适应MCMC采样方法。本书是视觉目标跟踪与MCMC采样方面的专著,反映了作者近年来在这一领域的主要研究成果。

本书可作为大专院校及科研院所图像处理、计算机视觉和智能视频分析领域的高年级本科生、研究生的教学和参考用书,也可供相关领域的教师、科研人员及工程技术人员作参考。

图书在版编目（CIP）数据

视觉跟踪中的马氏链蒙特卡洛方法 / 周修庄，鲁继文著．-- 北京 ：北京邮电大学出版社，2018.6

ISBN 978-7-5635-5388-4

Ⅰ.①视… Ⅱ.①周… ②鲁 Ⅲ.①蒙特卡罗法 Ⅳ.①O242.28

中国版本图书馆 CIP 数据核字（2018）第 023471 号

书　　　名:视觉跟踪中的马氏链蒙特卡洛方法
著作责任者:周修庄　鲁继文　著
责 任 编 辑:满志文　穆菁菁
出 版 发 行:北京邮电大学出版社
社　　　址:北京市海淀区西土城路10号(邮编:100876)
发　行　部:电话:010-62282185　传真:010-62283578
E-mail:publish@bupt.edu.cn
经　　　销:各地新华书店
印　　　刷:
开　　　本:720 mm×1 000 mm　1/16
印　　　张:7
字　　　数:120千字
版　　　次:2018年6月第1版　2018年6月第1次印刷

ISBN 978-7-5635-5388-4　　　　定　价:25.00元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前　言

视觉跟踪是计算机视觉研究的一个基本和重要环节。近年来，随着相关学科领域理论和技术的发展，视觉跟踪研究在理论方法和实现技术方面得到了进一步的深入和发展，并且在智能交通、人机交互、视频监控和军事等领域得到广泛的应用。在实际应用中，视觉跟踪方法通常需要处理一些复杂的情况，包括：如何建模目标的运动变化，以有效地捕获目标运动的不确定性；如何设计有效的外观模型以适应外观的显著变化；如何设计稳健高效的视觉跟踪算法以满足实际的目标跟踪任务等等。运动极大不确定性条件下的目标跟踪是视觉跟踪的一个难点，基于运动连续性假设的传统贝叶斯滤波跟踪方法在解决这一问题时具有很大的局限性。本书系统地阐述自适应马氏链蒙特卡洛采样的理论方法和实现技术，以及在目标运动极大不确定性条件下的视觉跟踪方法。

本书共分 4 章，主要内容包括：视觉跟踪的意义、应用和主要方法概述；马氏链蒙特卡洛采样和自适应马氏链蒙特卡洛采样方法；序贯随机逼近蒙特卡洛跟踪方法；运动突变的视觉跟踪方法。

第 1 章简要介绍视觉跟踪的意义、应用和主要方法概述。简要分析视觉跟踪的理论框架和基本技术，包括目标表示、外观建模、状态预测与估计等技术环节及其代表性算法。

第 2 章介绍马氏链蒙特卡洛采样方法及其在计算机视觉中的应用。首先介绍蒙特卡洛采样的基本理论、方法和实现技术。在此基础上介绍两种自适应马氏链蒙特卡洛采样方法及其算法实现，并给出它们在目标轮廓的多边形近似中的应用。

第 3 章介绍序贯随机逼近蒙特卡洛跟踪方法。通过分析贝叶斯滤波框架下的粒子滤波和基于马氏链蒙特卡洛采样的粒子滤波，引入能够有效克服粒子退化问题的序贯随机逼近蒙特卡洛采样方法。

第 4 章介绍运动突变的视觉跟踪方法。针对目标快速运动、视频镜头切换、低帧率视频等运动极大不确定性的视觉跟踪问题，介绍一种将目标分布自适应和提议分布自适应相结合的强自适应马氏链蒙特卡洛采样方法。

本书第 1 章至第 4 章由周修庄、鲁继文编写；全书由周修庄、鲁继文统稿。

本书是作者近年来研究工作的基础上编写而成。成书之际由衷地感谢本书作者周修庄的博士导师陆耀教授,感谢他多年的悉心培养和指导。同时,向本书参考文献的所有作者,以及为本书出版付出辛勤劳动的编辑老师们表示感谢。

限于作者的水平,本书难免有缺点和不足之处,恳请批评指正。

周修庄于北京邮电大学
2018 年 1 月 20 日

作者简介

1. 周修庄，男，博士，北京邮电大学自动化系副教授，博士生导师。2011 年毕业于北京理工大学计算机学院计算机应用技术专业，获得工学博士学位。现任职于北京邮电大学自动化学院。研究领域为计算机视觉、模式识别与机器学习，主要研究兴趣包括：人脸分析、视觉跟踪、度量学习，深度学习等。主持参与了国家自然科学基金、教育部博士点基金、北京市自然科学基金等多个国家级和省部级项目的研发工作。在 IEEE TPAMI/TIP/TIFS 和 IEEE CVPR/ACM MM 等人工智能领域国际权威期刊和顶级会议发表论文 40 多篇，2 篇入选 ESI 全球高被引论文。担任国际期刊《Neurocomputing》编委，ICME/IJCB/VCIP 等多个国际学术会议的 Special Session Chair。中国计算机学会计算机视觉专委会委员、中国自动化学会模式识别与机器智能专委会委员、中国图形图像学会视觉大数据专委会委员。

2. 鲁继文，男，清华大学自动化系副教授，博士生导师，IEEE 高级会员，中组部青年千人计划入选者(2015 年)。主要研究方向为计算机视觉、模式识别和机器学习。具体研究内容包括深度学习、度量学习、人脸识别和视频分析等。2012 年获得新加坡南洋理工大学电气与电子工程学院博士学位。2011 年 3 月至 2015 年 10 月在美国伊利诺伊大学香槟分校新加坡高等研究院工作，担任研究科学家，2015 年 11 月加入清华大学工作至今。近年来在领域国际权威期刊和顶级会议上发表论文 180 余篇，其中 IEEE Transactions 论文 51 篇(含 PAMI 论文 8 篇)，ICCV/CVPR/ECCV/NIPS 论文 27 篇，ESI 高被引论文和热点论文 6 篇，SCI 他引 1200 余次，Google Scholar 引用 4200 余次，获国际会议最佳论文奖或最佳论文提名奖 3 次。担任 IEEE 信号处理学会多媒体信号处理技术委员会委员，IEEE 信号处理学会信息取证与安全技术委员会委员，IEEE 电路与系统学会多媒体系统与应用技术委员会委员，Pattern Recognition、Journal of Visual Communications and Image Representation 等国际期刊编委，中国工程院院刊《Engineering》青年通讯专家，ACCV、ICIP、ICME 等 20 多个国际会议的领域主席或专题主席。

目　　录

第 1 章

绪　　论

1.1　视觉跟踪概述

据研究，人类通过感官感知外部世界的信息超过 70% 来自视觉。计算机视觉是在对人类视觉特性研究的基础上，利用计算机成像装置替代人类视觉器官作为信息输入手段，由计算机代替人脑完成对输入信息的处理和分析，使得计算机能够像人一样感知和理解外部世界，甚至完成人类难以完成的特定任务。计算机视觉是一门交叉性学科，涉及计算机科学、统计学、认知心理学、物理学、信号处理和分析，等等。近年来，随着各相关学科技术的发展，计算机视觉研究取得了令人瞩目的成就，在各个领域得到了广泛的应用。计算机视觉的研究起始于 20 世纪中叶。这个时期的研究工作主要是针对简单的二维图像进行识别和分析。在 20 世纪 70 年代，麻省理工学院人工智能实验室的 David Marr 教授提出了计算机视觉计算理论，该理论成为计算机视觉研究领域中第一个具有完整体系的理论框架，是计算机视觉研究领域一个里程碑式的工作，它极大地促进了计算机视觉理论的研究和发展[1]。

广义上，视频目标跟踪（视觉跟踪）是指对视觉场景中的运动目标进行检测、识别和跟踪，获得目标的运动、状态以及外观特性等参数，进而对跟踪的运动目标作进一步的处理与分析，实现对运动目标的更高层次的理解。视觉跟踪是计算机视觉领域中的一个研究方向。由于它具有极为广阔的应用前景，因而近年来已成为计算机视觉领域中备受瞩目的研究热点。视觉跟踪在军事国防、人机交互、智能监控、机器人导航、智能交通、生物医学工程等领域具有广阔的应用前景[2]。

在军事国防领域，视觉跟踪技术已应用于战场和基地侦查、导弹制导、地面和空中目标攻击等系统中，包括：机载系统对地面目标的监视以及地对空系统对空中目标的跟踪等；跟踪和分析试验飞机和导弹的轨迹和姿态；等等。在人机交互方面，传统的输入方式已不能很好地满足人与计算机的更为复杂的交流需要。视觉跟踪技术是计算机识别和理解人的各种姿态和动作、实现更为自然的人机交流的关键环节之一。在智能监控领域，系统需要对摄像机采集到的视频数据进行实时处理和分析，实现自动的目标检测、识别和跟踪。在此基础上对目标的运动轨迹进行分析。目标的检测、识别、跟踪和行为分析都是计算机视觉领域的主要研究课题，而有效的视觉跟踪技术是智能监控系统中核心的技术环节。智能机器人导航主要通过视觉信息识别周围场景中的物体，实现自主运动。在机器人手眼应用中，视觉跟踪技术可用于跟踪机器人所拍摄的场景中的物体、计算其运动轨迹，并根据需要选取最佳姿势完成特定的机械操作。智能交通系统是目前交通运输领域的研究热点。车辆的实时检测、跟踪和行人行为分析是智能交通系统的关键技术。系统通过对视频序列图像的自动分析，完成车辆和行人的检测与跟踪，进一步对车辆和行人的行为给出语义描述。其中需要解决的关键问题就是对车辆和行人的分割与跟踪。在生物医学工程领域，视觉跟踪技术在超声图像和显微图像的自动分析中有广阔的应用前景。例如，显微图像序列中的杂质与噪声严重影响了对图像中运动粒子（细胞）的提取和分析，视觉跟踪技术利用图像序列中目标运动和外观变化的连续性信息可以实现对感兴趣目标的跟踪与运动分析。

从 20 世纪 80 年代至今近 30 年的时间里，视觉跟踪的研究进展迅速，出现了众多的视觉跟踪方法。然而，由于视觉场景的高度复杂性，包括：目标运动不确定性、阴影、光照变化、背景噪声、目标干扰、遮挡、目标变形等，使得实现一个稳健、准确和快速的视觉跟踪算法依然非常困难[3]。传统的视觉跟踪技术一般都基于目标运动和外观变化较小的光滑性假设，因而在某些复杂场景中很难实现有效的目标跟踪。这些复杂的视觉场景包括如下几个方面。

1. 突变的动力学特性

随着视觉跟踪技术的应用领域的不断扩大，需要跟踪的运动目标的动力学特征也呈现出复杂多样化。在实际应用的复杂场景中，运动目标经常做不连续运动，包括自身不确定性的动力学特征和在外力情况下的突变运动[13]。例如，网球运动员的突然变速启动、乒乓球受球拍撞击后的变向运动。这种极大的运动不确定性破坏了传统跟踪算法的运动连续性假设，从而容易导致跟踪失败。



2. 低帧率视频

为节省存储代价，降低对网络带宽的过度依赖，在视频监控和互联网实时应用领域，对低帧率视频序列中目标的在线实时跟踪显得尤为重要。与此同时，低帧率往往导致目标运动的不连续和外观的显著变化。尽管基于随机采样的视觉跟踪算法，例如，粒子滤波[4]和基于马氏链蒙特卡洛的粒子滤波[5]等方法，能够对目标状态的不确定性进行有效的建模并维持多模假设，然而，目标的突变运动和外观的显著变化使得目前的大多数跟踪算法难以实现稳健、准确、快速的目标跟踪。

3. 多视角场景中的镜头切换

在多视角多摄像机的视觉场景中，由于摄像机镜头的切换导致目标在空间位置、尺度和外观方面的不连续变化。这种不连续变化很容易导致传统的跟踪算法失效。现有的大部分基于随机采样的概率跟踪算法为了实现精确的跟踪任务，需要使用大量的粒子在庞大的状态空间中逼近复杂的多模分布。因而，其计算复杂度非常高，难以满足实际应用中的跟踪需要。

4. 目标外观的显著变化

运动的极大不确定性，可能导致目标外观的显著变化。当前视觉跟踪领域的一个研究热点是基于外观模型(Appearance Model)的自适应跟踪。传统的视觉跟踪方法采用了静态外观模型，这种模型或者被人为地定义，或者是基于第一帧图像的训练样本，但是这些方法在所跟踪的目标发生外观变化时就容易导致跟踪失败。在复杂的现实场景中，一个自适应的外观模型是获得稳健的跟踪性能的关键。近年来，尽管基于生成模型(Generative Model)和判别模型(Discriminative Model)的视觉跟踪算法在目标外观发生变化时能够比较有效地实现自适应的跟踪，但是当目标外观发生显著变化时，如何更加有效地实现稳健的视觉跟踪至今仍是一项富有挑战性的研究课题。

作为计算机视觉领域的经典的、核心的研究内容之一，视觉跟踪技术融合了图像处理与理解、模式识别、机器学习以及计算机软硬件等众多学科领域的方法与技术。计算机视觉研究具有较高的科学价值和广阔的应用前景，而视觉跟踪是计算机视觉研究领域的基础和关键环节之一，是该领域中更高层次(如行为分析、理解等)研究的基础。综上所述，对于视觉跟踪技术的研究，来源于计算机视觉应用领域的实际需求。为了解决这一课题，需要进一步探索更为有效的视觉跟踪方法和技术。因此，本课题的研究具有重要的学术和应用价值。

1.2 国内外研究现状及发展趋势

视觉跟踪在计算机视觉的相关领域得到了广泛的研究和应用。在智能视频监控领域，伯克利的视觉研究组在20世纪90年代开展了初步的研究工作，提出了多个有效的、能够初步处理遮挡问题的车辆跟踪算法。同时期，在智能机器人研究领域，卡耐基梅隆大学机器人研究所提出了智能控制和视觉技术相结合的视觉跟踪方法。在视频分析领域，麻省理工学院媒体实验室提出了体育视频中多目标跟踪方法。在智能人机交互领域，伊利诺伊大学香槟分校在目标识别、行为分析和视觉跟踪方面开展了一些卓有成效的工作。近年来，澳大利亚格拉斯理工大学的计算机图形与视觉研究所，以及苏黎世联邦高等工业大学的视觉研究组在视觉跟踪领域开展了基于在线学习外观模型的研究工作，并取得了一些具有重要影响的成果。此外，在国内，一些高校和研究机构近年来也在视觉跟踪领域陆续开展了这方面的研究工作，这些研究机构包括中科院自动化所、中科院计算所、清华大学、北京大学、西安交通大学、哈尔滨工业大学和华中科技大学等。

近年来，视觉跟踪问题已成为计算机视觉领域的研究热点之一，一些重要的国际期刊和国际学术会议，例如：IEEE Transactions on PAMI，IJCV，CVPR，ICCV，ECCV等，发表了大量有关视觉跟踪方面的研究论文。其中一些学术会议就视觉跟踪问题组织了专门的工作组进行技术讨论。如前所述，视觉跟踪之所以引起如此广泛的关注，主要是由于它能够广泛应用于社会生产生活的诸多领域。

视觉跟踪的研究内容涉及目标表示、特征选择、外观建模、目标检测识别、滤波技术、最优化方法和随机采样方法等众多相关领域的理论方法和实现技术，具有较高的研究难度和挑战性。本节首先对视觉跟踪的基本方法和技术进行简要的分类与总结，然后对视觉跟踪领域的主要研究工作进行回顾与评述。

1.2.1 目标表示、特征选择与外观建模

在视觉场景中，被跟踪目标的形状和外观建模是视觉跟踪的基础与关键。表示模型的优劣对跟踪算法的性能有重要影响。目标的形状表示方法可以大致归纳为如下几种[6]。

(1) 使用单个点[7]或一组点[8]表示目标。当被跟踪目标在图像中占有较小区域的时候，这种表示方法虽然简单，但通常也比较有效。



（2）使用矩形框或椭圆[9]表示目标。这种表示方法在视觉跟踪研究中最为常见，其特点是比较适合目标的仿射变换与射影变换，可以用来近似地表示刚性目标与非刚性目标。

（3）使用轮廓[10]表示目标。轮廓通常应用在复杂非刚性目标的视觉跟踪，在这种情况下简单的几何形状难以有效地表示目标的轮廓。

（4）基于部件或骨架的目标表示[11]。具有关节特性的目标可以认为它是由关节或部件链接在一起的，这种表示方法便于结合特定的运动模型，对具有关节体的运动目标进行精确的跟踪。此外，骨架模型可以用来表示刚性和非刚性目标，它通常通过图像变换从目标轮廓内部区域中提取而得到。

给定目标表示方式选择恰当的图像特征来刻画被跟踪目标，在视觉跟踪中具有重要的作用。特征选择通常是面向问题的，“好”的特征通常应具备较好的区分性或判别力。特征的选择通常依赖于特定的目标表示方法。例如，颜色用来作为基于直方图的外观表示的特征。而对于基于轮廓的目标表示方法，边缘通常用来作为相应的特征。通常，跟踪算法使用若干特征的组合。视觉跟踪中通常采用的基本的特征如下[6]。

1. 目标的颜色

在计算机视觉领域，RGB 颜色空间是最为常用的，此外还有 Lab 颜色空间和 HSV 颜色空间也较为常用。这三种颜色空间的差别在于：RGB 对人眼的色彩感知上是非均匀分布的，各个分量之间并不独立；相比 RGB 空间，Lab 颜色空间是感知均匀的，而 HSV 颜色空间是感知近似均匀的。总而言之，这几种颜色空间在视觉跟踪领域的应用都在一定程度上依赖于特定的问题和任务，并不存在一种颜色空间对所有问题是普遍适用的。

2. 目标的边缘

在图像处理领域，通常利用目标在边界处的强度的显著变化来检测目标的边缘[12]。边缘特征一般对场景光照变化不敏感，这个特性对于视觉跟踪来说非常重要。常用的边缘检测算子包括 Canny 算子、Sobel 算子和 Laplacian 算子等。

3. 目标的纹理

一般认为，纹理是图像局部亮度变化特性的一种描述，这种特性包括规则性或平滑性[1]。与颜色特征不同，纹理特征对于场景光照变化不敏感。常用的纹理描述有多种，包括：灰度共生矩阵、Laws 纹理描述子和小波等。

4. 光流场

光流特征在计算机视觉以及图像处理中有着广泛的应用，它与运动检测和运

动估计密切相关。在物理中，运动可以用运动场描述；而在一个图像平面上，目标的运动往往是通过图像序列中不同像素点的灰度强度分布的变化来体现的。空间中的运动场变换到图像平面上就成为了光流场。计算光流的经典方法包括：Lucas-Kanade 方法[14]、Horn-Schunck 方法[15]等。

此外，在视觉跟踪领域常用的目标特征还包括：梯度方向直方图（Histogram of Oriented Gradient, HOG）[16]，局部二元模式（Local Binary Pattern, LBP）[17]，尺度不变特征变换（Scale-Invariant Feature Transform, SIFT）[18]等。一般情况下，用于视觉跟踪的特征的选择是依赖于特定跟踪任务的，不同的跟踪问题需要人为地选择合适的特征或特征组合。近年来，自动的特征选择在视觉跟踪领域得到了广泛的研究[19,20]。基于 AdaBoost 方法的在线特征选择是此类方法的典型代表。

基于目标表示方式和选定的特征（集合），用于视觉跟踪的目标外观模型包括如下几种。

1. 模板

模板（Template）是视觉跟踪中最为常用的外观表示方法。它一般使用简单的几何形状（如椭圆、矩形等）来表示[21]。特定设计的模板，例如基于分块表示的模板[22]，可以同时包含目标的空间和外观（颜色等特征）信息。

2. 概率分布

目标外观的概率分布是另一种较为常用的外观表示方式。参数化的概率分布表示可以采用单个高斯分布[23]或混合高斯分布[24]；而非参数的表示方式主要是核密度估计[25,9]。

3. 主动外观模型

在主动外观模型（Active Appearance Model, AAM）[11]中，目标的形状由一组特征点来描述，这组特征点位于目标边界或目标区域内部。每一个特征点对应于一个外观特征（颜色、边缘、纹理等）向量。AAM 主要特点在于它同时把目标的形状和外观特征模型化。

4. 多视外观模型

在复杂的场景中，当运动目标的外观变化较大且遮挡频繁发生时，采用多摄像机的多视角（multi-view）跟踪是个不错的选择。与此相对应的多视外观模型，用来模型化目标的不同视角。多视外观模型通常考虑从不同视角的样本中生成一个子空间[26]，或者考虑从不同视角的样本中训练得到一个判别分类器[27]。



1.2.2 视觉跟踪方法

前面的章节已经对目标表示和外观建模做了简要的归类和介绍。概括地说，视觉跟踪问题可以分解为：目标表示、外观建模和跟踪框架（方法）。在视觉跟踪研究领域，外观建模通常分为：前景建模和前景背景联合建模。与此相对应的外观模型可以归结为两大类：原生模型和判别模型。原生模型试图构建一个紧凑的外观模型来描述可能的目标外观变化。判别模型不同于原生模型，旨在将目标外观与背景区分开来。大多数判别外观模型是通过机器学习的方法学习一个两类分类器（Binary Classifier）来区分跟踪的目标和背景。至于跟踪框架（方法），大致可以归类为：确定性的跟踪方法和基于采样的随机跟踪方法两种。在视觉跟踪应用中，跟踪方法通常与外观建模方式有关。在这一节里，我们对视觉跟踪的主要方法做一个简要的回顾与总结。

1. 原生外观模型

1）基于目标模板

目标模板既可反映目标的空间信息，也可包含目标特征（颜色、纹理等）信息。在基于模板的视觉跟踪研究方面，较早期的工作是 1981 年 Lucas 和 Kanade[14] 的图像配准方法。这一开创性研究工作的基本思路是：将首帧图像的目标区域（矩形）设定为模板，在后续的图像序列帧中搜寻与设定模板匹配度最高的区域。采用的匹配准则是目标模板和候选矩形区域对应像素灰度值差的平方和（Sum of Squared Differences，SSD）。在文献[28]中，只能获得平移运动参数的 Lucas－Kanade 方法被扩展到可以获得仿射变换参数，并应用于特征点跟踪。在这一研究工作中，作者还提出了在评估特征点质量的方法：计算当前帧特征点与首帧特征点的不相似度，当不相似度大于一定阈值时认为该特征点不再可靠进而被丢弃。Shi 和 Tomasi[28] 认为，一个“好”的特征点的选择取决于这一特征点是否能给跟踪算法带来更好的跟踪性能。此外，Baker 和 Matthews[29] 对经典的 Lucas-Kanade 方法做了全面的回顾，并对该框架下的各种算法进行了较为全面的评述。

视觉跟踪过程中使用静态不变的外观模板是不可行的，这是因为目标外观在视觉场景中受各种内部或外在因素影响而不断变化。因而，为了获得稳健的跟踪性能，模板更新是一个必须考虑的重要问题。一种非常简单的更新策略是使用当前跟踪的结果对目标模板进行更新。这一更新策略的问题在于跟踪过程中误差累积将最终导致目标漂移（Drifting）并导致跟踪失败。针对这一问题，Matthews 等人[30] 给出了一种解决模板更新问题的方案：在跟踪过程中保留首帧的模板，用当

前帧的跟踪结果对目标模板进行更新，并和首帧模板进行对齐。在文献[31]中，作者提出用一个参数模型来建模目标的运动，用一个子空间来建模场景中目标光照的变化，用稳健估计方法来建模目标的遮挡。此外，Nguyen 和 Smeulders[32]针对遮挡情况下的模板匹配问题进行了深入的研究，并提出对目标模板中的每一像素分别调用卡尔曼滤波器进行更新。该方法对于场景中目标光照变化具有较好的稳健性。在文献[33]中，Jepson 等人提出了一种可以处理光照变化和外观变形的跟踪方法。它融合目标稳定部分、帧间差别部分和噪声影响部分的 WSL 三分量模型。该方法使用小波作为目标特征，采用基于在线的 EM 算法更新模型参数。在这一模型中，目标的仿射运动估计是通过目标稳定分量进行。该模型的主要缺陷在于稳定分量部分，当背景变化较缓慢时，部分背景也将被估计成目标的一部分。同样基于在线估计参数的方式，Han 和 Davis[34]提出使用高斯混合模型对目标区域内的像素进行在线建模。在该模型中，作者通过自适应带宽的 Mean Shift 算法在线自动估计各个模型参数。使用单一目标模板的跟踪方法在实际跟踪任务中应用广泛，但在目标姿态变化较大时往往导致跟踪失败。

2）基于目标特征的概率分布

这种目标外观建模方法通常采用目标内部的特征（颜色、边缘等）的概率分布来描述目标。这种建模方法的主要特点在于它忽略了目标区域的空域信息，因而能用于处理目标的非刚性形变。这方面的早期研究工作包括：Birchfield[35]采用一个椭圆表示人脸并使用颜色直方图对人脸进行跟踪。同样基于颜色直方图表示，Comaniciu 等人[36]结合核密度估计方法，在计算颜色直方图时使用核函数对直方图进行加权使得越靠近目标中心的像素权重越大，从而在一定程度上避免了嘈杂背景的干扰。此外，作者还提出了处理目标尺度变化的方法，它通过对当前跟踪得到的尺度进行一定范围的伸缩（例如，5%，10%等），然后分别计算相似度，相似度最大的作为最终的目标尺度。Comaniciu 等人的方法为目标（位置和尺度）跟踪提供了一种新的思路。该方法的一个缺陷是基于核函数的颜色直方图并不能很好地描述目标的尺度，往往在匹配时小尺度比大尺度具有更高的相似度，从而造成尺度跟踪漂移现象。之后，Collins[37]提出了一个基于尺度空间的处理方法。该方法采用一个尺度核函数，在空间位移和尺度变化上交替迭代调用 Mean Shift 算法，从而获得目标的空间位置和尺度。

在使用颜色直方图方面，Mean Shift 等确定性跟踪方法在多模复杂场景下往往易于陷入局部模式从而导致跟踪失败。针对这一问题，在文献[38]和[39]中提出了结合随机跟踪方法和颜色直方图的处理方案。在文献[38]中，Pérez 等人提出



在粒子滤波框架下基于颜色直方图的跟踪方法,该方法对复杂多模场景具有一定的稳健性。当使用大量粒子时,对颜色直方图的计算是个很大的计算负担。为了减少计算代价,Porikli[40]提出使用积分直方图。与此类似,Adam 等人[41]提出使用基于分块目标表示的颜色直方图,其中每一个子块都用颜色直方图来表示。基于这种外观模型的跟踪方法对局部遮挡具有较好的稳健性。此外,文献[42]和[46]还给出了处理更多运动参数(包括位移、尺度、旋转方向等)的解决方案。这些方案大多基于椭圆表示的颜色直方图模型,采用核函数进行跟踪。

对于非刚性形变的目标跟踪,基于目标特征的概率分布的跟踪方法能够获得较好的跟踪性能,具有一定的稳健性。不过,当场景中具有多个相似特征的目标,或在嘈杂的背景下,这种基于概率分布的外观模型极易导致跟踪漂移和失败,其根本原因在于这一外观建模方式忽略了目标的空间分布信息。

3）基于子空间

基于子空间的跟踪方法通常是通过训练样本学习(在线或离线)构造(通常通过主成分分析)一个子空间模型。该方法的主要优点在于它同时兼顾了目标的光照变化与姿态变化,这使得它能够处理较为复杂场景下的视觉跟踪任务。

基于离线构造子空间方面,Black 和 Jepson[26]给出了多视角子空间构造方法。值得注意的是,该方法通过深入分析光流计算中的灰度强度不变假设,提出了子空间的一致性假设。作者认为,多视角目标观测可以采用相同的特征基(集)进行重建,使得重建图像与原始图像具有一致的灰度强度分布。基于这一假设可以实现不依赖于特定视角的视觉跟踪任务。

确实,基于离线的构造方法尽管也能在一定程度上描述和适应目标的外观变化,但是基于离线的方式在复杂的场景中往往由于累积的跟踪误差而导致跟踪漂移。针对这一问题,Ross 等人[47]提出了基于在线学习渐进构造子空间基的算法。在该算法中,子空间基是随着顺序获得的外观样本的逐渐增加而不断提升其准确性的。作者的实验表明,该算法在处理显著的光照、姿态和尺度变化等方面具有较强的稳健性。此外,在 Log 欧氏黎曼度量之下,Li 等人[48]提出了一种基于 SLAM 模型的在线跟踪算法来处理目标姿势、尺度自适应变化的情况。Wu 等人[49]基于渐进的协方差张量学习提出了一种姿势和外观自适应的实时跟踪算法。Li 等人[50]通过自适应地更新样本均值和特征基,渐进地学习一个低阶张量特征空间表示来建模目标的外观模型,提出了一种在线的张量子空间学习算法来实现有效的目标跟踪。Yang 等人[54]提出一种基于子空间的外观自适应策略,它基于三个约束,即负样本数据、自底向上的成对正负样本数据以及自适应的目标动力学特征。

尽管基于子空间的跟踪方法取得了较好的跟踪性能，但是在复杂场景下构造和更新子空间模型以实现目标的自适应变化仍然是个极富挑战性的问题。

2. 判别外观模型

相比原生外观模型只对目标自身进行建模，判别外观模型同时对目标和背景进行联合建模，因而有可能获得较好的区分目标与背景的判别能力。近年来，基于判别外观模型的视觉跟踪受到了国内外研究人员的广泛关注。在这类方法中，视觉跟踪被看作一个分类的问题。换句话说，在每一帧中跟踪器的任务就是以最佳方式区分目标与背景。

Avidan[27]使用支持向量机(SVM)离线训练一个分类器用于对目标与背景的分类，实现了一个有监督的目标跟踪器。该方法的优点是：在跟踪阶段由于不用在线训练分类器因而跟踪速度较快。缺点是由于分类器是事先就训练获得的，因而被跟踪目标必须是事先就已经知道的；此外，在跟踪过程中由于分类器是固定不变的，因而它很难适应不断变化的背景。为了解决这一缺陷，Collins 和 Liu[51]提出一种基于在线特征选择的自适应跟踪方法。作者使用 RGB 的线性组合为特征，通过对前景和背景点的采样获得它们对于不同特征组合的分布情况，选择最有判别性的特征组合生成置信图，最后采用 Mean Shift 算法进行模式定位。与此类似的工作包括：Yin 和 Collins[52]在文献[51]工作的基础上，提出对目标的背景进行分割，然后对每对目标和背景使用不同的特征，然后对生成的置信图使用 Mean Shift 算法进行目标定位。Wang 等人[53]使用粒子滤波框架自动进行特征选择，并在粒子滤波框架下实现目标的随机跟踪。

一个判别模型的重要工作是 Avidan[55]提出的集成学习(Ensemble Learning)框架下的跟踪方法，作者采用 AdaBoost 算法在线地训练获得前景/背景两类分类器，并用获得的分类器对后续帧进行分类并获得一张置信图，最后运用 Mean Shift 算法进行目标定位。作者也讨论了如何在线更新弱分类器以训练强分类器的方案，以适应目标外观与背景的自适应变化。Grabner 和 Bischof[56]设计了一种基于在线 Boosting 的特征选择算法实现实时的目标跟踪。这类基于 Boosting 的判别外观模型可以有效地将被跟踪目标与背景区分开来，然而它最大的缺点是因跟踪过程中跟踪误差的不断累积，从而可能导致跟踪漂移。为了解决跟踪漂移问题，Tang 等人[57]提出一种基于联合训练(co-training)框架在不同的特征集上训练多个 SVM 分类器来提高外观的判别能力；同样基于联合训练框架，Liu 等人[58]通过寻找 Boosting 错误上界使用多视弱假设在线地更新外观模型。Woodley 等人[59]提出一种结合原生模型和判别模型的跟踪算法。尽管这些算法能够在一定程度上



缓解跟踪漂移问题，但是在复杂场景中当目标外观显著变化时，这一困难并没有得到很好的解决。

值得注意的是，当更新自适应外观模型时，如何选择正负样本构建训练集。通常的策略是，将当前跟踪器估计的目标位置作为正样本，而从目标周围的领域中选取负样本。然而，如果跟踪器估计的目标位置是不准确的，那么外观模型将会以这个不准确的正样本对模型进行更新。随着跟踪的进行，这种方法会使模型的判别性能严重下降，并且将可能导致跟踪漂移。

如果使用固定的判别分类器进行目标跟踪，那么它将会有效地避免跟踪漂移问题，但是它缺乏必要的外观自适应能力。同样地，如果使用在线的自适应外观模型，那么可能会导致跟踪漂移问题。这似乎是个两难的问题，因而，如何有效地避免跟踪漂移问题同时还能保留在线的自适应能力，已成为近年来外观模型研究的热点方向。

最近，Grabner 等人[60]将半监督学习引入到视觉跟踪问题，基于 SemiBoost[61]算法，他们提出了一个基于半监督学习的在线分类器实现稳健的目标跟踪。该方法从先验信息（例如，使用第一帧中初始的目标状态）获得标记样本（Labeled Sample），而未标记样本（Unlabeled Sample）则是从跟踪过程中获得的。基于半监督学习技术，从跟踪过程获得的未标记样本集能够帮助跟踪器学习一个在线自适应的外观模型，而仅从先验信息中构建标记样本集则能够有效地克服跟踪漂移问题。

在基于半监督学习的判别外观模型中[60]，标记样本只来自于初始的第一帧图像，而后续训练样本不被标记。这种方法在目标完全离开视觉场景，或者出现遮挡时表现出良好的判别性能，但是在帧间运动较小的情况下，一些有用的信息，例如运动连续性信息并没有得到很好的利用。Barbenko 等人[62]提出了基于多示例学习（Multiple-Instance Learning，MIL）的外观模型更新算法实现对目标的稳健跟踪。事实上，基于 MIL 学习的跟踪方法受到 Viola 等人[63]提出的基于 MIL 学习的目标检测方法的启发。通常，目标检测存在一些固有的歧义性，这种歧义性使得仅仅通过传统方法训练分类器显得困难重重。基于这个原因，Viola 等人[63]提出了运用多实例学习方法进行目标检测。在 MIL 的训练过程中，实例是以集合，或称为“包（bag）”的形式呈现，并且标记是以包而不是个体实例而体现。如果一个包被标记为正的，那么认为这个包至少包含一个正实例，否则这个包就被认为是负的。较之目标检测，目标跟踪具有更多的歧义性，这是由于跟踪器没有人为的干预输入同时需要自我引导。因而从理论上说，将 MIL 技术应用于目标跟踪问题比将其用于目标检测问题更为适合。此外，为了实现一个更为稳健的跟踪算法，Zeisl

等人[64]基于半监督 Boosting 的方法设计了一个在线 MIL 跟踪算法。他们在多种具有挑战性的跟踪测试集上进行的实验表明,与当前同类算法相比,该算法可以获得更加稳健和稳定的视觉跟踪。

3. 基于确定性的跟踪方法

从方法论的角度看,视觉跟踪的研究方法可以归结为自底向上(bottom-up)和自顶向下(top-down)的两种思路。自底向上的方法通常不依赖于先验知识,而是通过可靠的检测(Detection)直接从视觉场景中获取目标的相关信息并用于视觉跟踪。自顶向下的方法也称为模型驱动(Model Driven)的方法,最典型的例子是线性或非线性的状态空间(State Space)模型的方法,它通常是在贝叶斯框架下,通过给定的先验分布和观测模型,迭代的求解后验分布进而实现目标状态的估计。此外,也有一些研究工作将这两种研究思路相结合。对应于自底向上和自顶向下的两种研究思路,现有的视觉跟踪算法大抵可以分为两大类:确定性跟踪方法和基于采样的随机跟踪方法。

确定性跟踪方法一般利用自底向上的底层视觉信息提供的可靠检测结果,并使用迭代的优化技术[65]或线性规划方法[78]实现快速的目标跟踪。这些自底向上的视觉信息包括前面章节所述的目标特征,即 Haar-like 特征、颜色直方图、梯度方向直方图(HOG)、边缘方向直方图(EOG)、尺度不变的特征变换(SIFT)等。Mean Shift[65,66]技术被广泛地应用于视觉跟踪领域[36,37],是确定性跟踪方法的典型代表。作为一种基于非参数优化的跟踪算法,它独立于目标特征且具有收敛速度快、计算复杂度低等优点。Cheng[65]阐明了 Mean Shift 本质上是个自适应步长的梯度下降算法。Fashing 和 Tomasi[67]进一步指出 Mean Shift 和 Newton-Raphson 优化算法的内在关联。自从 Comaniciu 等人[36]首次将 Mean Shift 引入到视觉跟踪领域,至今学者们提出了许多改进的 Mean Shift 跟踪算法,包括:Elgammal 等人[25]提出在目标区域中融合空间结构信息的 Mean Shift 跟踪算法,克服了忽略空间分布信息的颜色直方图跟踪方式的缺陷;在文献[43]中,作者基于多个空域分布核函数提出了一种能精确地估计目标尺度和方向的跟踪算法;Fan 等人[46]给出一个基于核跟踪的相似度度量的理论分析,并探讨了避免在核跟踪算法中产生奇异性的核函数设计准则与基本策略。尽管基于 Mean Shift 的确定性跟踪方法取得了很大的成功,但是这种方法通常要求相继的两帧之间目标的位移较小。其主要原因在于 Mean Shift 算法容易受到多模场景中背景噪声等因素干扰而陷入局部最优状态,从而导致跟踪失败。

鉴于 Mean Shift 跟踪算法的局部优化本质,一些学者提出结合全局采样与局



部优化的策略实现对快速运动目标的跟踪。Shan 等人[68]通过使用颜色和运动信息，将 Mean Shift 局部优化过程嵌入到粒子滤波框架实现对手的实时跟踪，作者的实验表明这种结合大大降低了随机跟踪所需的粒子数目，并从理论上分析了如何确保在使用局部优化之后粒子集仍能近似地服从真实后验分布。基于类似的思路，Shen 等人[69]提出在多带宽 Mean Shift 跟踪框架下引入模拟退火思想，渐进地平滑化代价（能量）函数表面，从而最终获得全局最佳模式。使用一个大带宽过度平滑化核密度函数通常使得密度分布具有较为简单的单模（Unimodal）特性。在退火过程中，缓慢地逐渐减小带宽，并调用 Mean Shift 过程迭代地最小化代价函数。理论上，只要这个带宽减小过程足够缓慢，算法最终将会收敛到全局最优模式。

对于运动具有极大不确定性的目标跟踪问题，确定性的跟踪方法容易跟踪失败并且很难从错误状态中恢复。尽管结合随机采样或模拟退火过程的 Mean Shift 算法能够在一定程度上提高跟踪的稳健性，但在处理复杂多模分布的问题时仍然难以取得令人满意的跟踪性能。

4. 基于随机采样的跟踪方法

不同于确定性的跟踪方法，基于随机采样的概率跟踪方法能有效地实现在多模分布场景下的视觉跟踪任务。这类方法通常是在基于贝叶斯跟踪框架下，通过状态预测与观测似然迭代地估计后验分布实现对目标状态的估计。在运动模型与观测度量呈线性，噪声服从高斯分布的情况下，卡尔曼滤波[70]可以获得贝叶斯滤波的精确解析解。对于非高斯条件下的非线性系统，通常是使用蒙特卡洛模拟技术来实现迭代的贝叶斯滤波。粒子滤波[4]是此类跟踪方法的典型代表。它通过一个加权的粒子集近似地表示滤波分布，其中粒子是从滤波分布的状态空间中随机采样的，后验的更新则依赖于粒子的序贯传播，粒子权重通过观测似然进行更新。在实际的粒子滤波应用中，通常需要考虑跟踪精确性与计算代价（所需粒子数目）的平衡。一般情况下，使用大量的粒子能够很好地近似复杂的滤波分布，但是高昂的计算代价往往使跟踪算法不能满足实际的跟踪需要。

一般来说，直接从后验（滤波）分布采样是比较困难的。在状态空间中有效的采样策略对于粒子滤波算法来说至关重要。从这个角度看，粒子滤波算法可以分为：重要性采样[71~74]，因子采样[4]，以及拒绝采样[75]等。重要性采样的基本出发点在于尽可能从样本空间中靠近后验模式的区域采样。为了达到这个目的，重要性采样方法通常需要使用一个易于采样，而且其分布靠近后验模式的提议分布（Proposal Distribution）。对于贝叶斯滤波问题，重要性采样的关键在于如何选取提议

分布。序贯重要性采样(Sequential Importance Sampling, SIS)是最为常用的一种解决贝叶斯滤波问题的算法,它通过先前时间阶(Time Step)估计得到的后验密度迭代的产生提议分布。序贯重要性采样算法的缺点在于粒子退化问题[71],也就是说,一旦提议分布与实际滤波分布的形状匹配不佳,大部分粒子的权重将变得很小,这些粒子在传播过程中其影响将逐渐趋于零,这将导致大量的计算代价浪费在低权重的粒子上。另一方面,因子采样算法的缺陷在于样本贫化(Sample Impoverishment)现象,即多个样本落在样本空间中的同一点,导致样本集缺乏多样性(Diversity)。

通过在状态空间中维持多个假设,粒子滤波算法能够比较有效地处理多模分布问题。作为一个一般性的通用跟踪框架,尽管它在视觉跟踪领域取得了很大成功,但是粒子滤波算法本身在处理高维状态空间的跟踪问题时仍然具有局限性(样本贫化等)。为了克服这一局限性,一些研究人员尝试通过多种途径来进一步改进粒子采样效率。近年来,将密度估计方法引入到粒子滤波框架成为了提高粒子采样效率的有效途径之一[76-77]。Cham 和 Rehg[79] 使用一个分片的高斯模型(Piecewise Gaussian)来表示观测似然度量和先验预测密度,从而获得后验密度的解析计算方式,实现多假设跟踪。Vermaak 等人[80] 则直接将后验密度建模为高斯混合模型,作者指出后验密度的传播可以通过对混合模型的单个成分(Component)独立地进行滤波,每个成分都获得一定数目的粒子,以此实现多假设跟踪。Han 等人[76]将核密度估计技术引入到序贯贝叶斯滤波框架实现快速的后验密度解析和估计。基于高斯混合模型假设,作者首先使用 UT 变换(Unscented Transform)计算先验密度;然后从观测数据中通过多阶段采样和密度插值的方式估计观测度量的解析形式;最后,将这两项估计值相乘并使用可变带宽的 Mean Shift 算法获得后验密度的光滑近似。此外,密度估计的支持向量(Support Vector)方法也在序贯贝叶斯滤波中得到了应用[81]。Baner jee 和 Burlina[81] 采用稀疏核密度估计的方法计算给定粒子集的先验密度,其中的密度估计器包括支持向量描述(SVDD)和相关向量机(RVM)[82]。作者使用 SVDD 算法把目标分布表示为一个线性混合核函数(LMK),结合序列的观测模型迭代地给出后验分布的闭合形式表示。在粒子滤波框架中采用支持向量的方法估计密度,其目的是获得后验密度的稀疏表示和解析解,从而降低粒子滤波的计算代价。

尽管基于密度估计的技术应用在贝叶斯滤波有助于简化采样问题,同时在采样过程中也可以维持多个假设,但是粒子滤波算法仍然受限于高昂的计算代价,尤其是在复杂滤波分布条件下需要使用大量的粒子来近似后验分布。为了降低计算



负担从而提高粒子采样效率，学者们提出了多种粒子滤波改进算法。Deutcher 等人[83]在传统的粒子滤波跟踪算法(Condensation)[4]中交错地使用一个模拟退火过程。模拟退火过程便于从平滑化的后验分布采样，这种结合使抽取的样本靠近实际的后验模式，从而避免采样陷入局部最优模式。通过在一个高维状态空间中关节体跟踪(Articulated Body Tracking)的实验，作者给出了该采样方法与 Condensation 算法的全面的性能比较分析。该方法的一个局限性是由于在采样中调用模拟退火算法，缓慢的退火过程使滤波算法的计算复杂度仍然很高。Choo 和 Fleet[84]提出在序贯贝叶斯框架下使用马尔科夫链蒙特卡洛(MCMC)采样方法替代广泛采用的重要性采样算法，直接从后验分布采样。这种采样方式主要是通过一个合适的提议分布(Proposal Distribution)从后验分布中采样，以此构建一条收敛于后验分布的马氏链。相比与传统的序贯重要性采样算法，使用 MCMC 采样的好处在于可以更高效地从后验分布中采样，因而可以使用较少的粒子来近似目标分布。与此工作类似，Khan 等人[5]提出一种基于 MCMC 采样的粒子滤波算法用于多目标跟踪。作者直接在高维的状态空间中建模多目标状态，通过引入一个基于马尔科夫场的运动先验来描述多目标跟踪场景下目标间的相互影响(Interaction)，实现在高维状态空间中的 MCMC 采样。此外，为了处理可变数目的多目标跟踪问题，作者提出使用可逆跳跃 MCMC(RJMCMC)建模可变维状态空间中的后验采样。他们的实验表明相比于独立的多个粒子滤波实现，使用 MCMC 采样替代序贯重要性采样可以显著地减少所需的粒子数目，从而降低计算复杂度。紧接着，Zhao 等人[85]提出采用数据驱动(data-driven)的 MCMC 采样思路替代传统的 MCMC 方法，在采样过程中更有效地估计提议分布，使产生的样本更靠近样本空间中的局部模式，从而提高采样效率、降低计算复杂度。

在视觉跟踪研究领域，已有的大多数跟踪算法都基于一个假设，即目标运动(状态变化)和外观变化的连续性。然而，在一些复杂的场景中，这种连续性假设并不符合实际情况。运动连续性假设对视觉跟踪至关重要，这是由于这个假设有利于跟踪算法可以在较小的状态空间中建模目标的运动，使在有限的计算资源条件下实现可靠的跟踪任务。一旦这种连续性假设不再成立，那么就需要在状态空间中扩大搜索(采样)范围来估计目标的状态，这通常需要增加大量的样本(粒子或迭代)来实现在庞大的方案空间中进行有效的采样。

在视觉跟踪领域，目标运动的不连续性(运动突变)主要包括：目标自身动力学特性的极大不确定性，外在因素影响的动力学改变，摄像机镜头切换，低帧率视频等。较早对目标跟踪中的运动动力学特性进行研究的是 Blake 等人[86,87]，他们从

特定场景的训练数据中学习动力学特征并用于动力学变化的状态建模。这种基于运动特征学习的方法对于具有一定运动模式的特定场景下的跟踪被认为是可行的。然而由于基于特定的训练集,这种方法很难适合于一般性目标的跟踪,因而可扩展性较差,难以应用于低帧率视频等复杂运动条件下的视觉跟踪任务。多尺度和分层采样的策略是处理目标运动不连续问题的另外一种方案。在贝叶斯框架下,Sullivan[88]等人结合不同尺度下的图像观测,提出在状态空间中进行有效搜索的跟踪方案。这种方法的缺点是,大尺度下的不准确的推断将导致在原始尺度下目标搜索的失败。针对这一困难,Hua 等人[13]提出在多尺度下的一种基于动态马尔科夫网络(Dynamic Markov Network)的合作搜索策略,并设计了一个序贯信念传播(Sequential Belief Propagation,SBP)算法以实现更为准确的贝叶斯推断。作者的实验表明,这种方法能够有效地处理多种运动不确定性条件下的视觉跟踪问题。

近年来,一些研究人员对运动突变的视觉跟踪问题进行了深入的分析和研究。针对低帧率条件下的视觉跟踪问题,Li 等人[89]提出使用一个级联的粒子滤波器(Cascaded Particle Filters)结合多个在线(或离线)学习到的检测模型来提高粒子采样效率和目标跟踪准确度。该方法在多个低帧率测试实例中取得较好的跟踪效果。然而,它需要一个事先训练好的可靠的观测模型和一个离线的学习过程,此外,它不能处理大范围快速运动的目标跟踪,例如多摄像机镜头切换等。最近,Kwon 和 Lee[90]提出了一种基于 Wang-Landau 采样[91]的能够处理多种运动突变的目标跟踪算法。它的基本思想是使用 Wang-Landau 算法在整个样本空间中对滤波分布进行全局采样从而避免陷入局部的最优模式。由于能有效地避免在采样过程中陷入局部最优状态,这种方法在多种不连续运动场景中取得了很好的跟踪结果。然而,该方法所使用的提议分布是基于均匀分布的全局跳跃(Global Stochasitc Hopping)模式,没有考虑局部模式的细节提炼(Detailed Mode Refinement),尽管作者使用了一个退火模式来分阶段压缩(Shrink)样本空间,整个采样过程仍然需要相当数目的粒子来近似复杂的滤波分布,因而在实际应用中,离实时跟踪需要还有一定的距离。Wang-Landau 算法源自早先在统计物理中有着广泛应用的 Multicanonical 采样算法[92],它在物理学中的应用主要包括态密度(Density of States, DoS)和其他一些重要物理量的计算。尽管 Wang-Landau 算法在统计物理学中的应用取得了很大的成功,但是至今还没有坚实的理论证明 Wang-Landau 算法在一般状态空间中的收敛性,并且在一些统计学和物理学应用中已发现它只能取得有限的统计精度[93]。



针对 Wang-Landau 采样算法在贝叶斯滤波跟踪中的采样效率问题，本文作者[94]在序贯贝叶斯滤波框架下，将随机逼近蒙特卡洛(SAMC)[93]采样方法引入到贝叶斯滤波框架中，提出了一种序贯随机逼近蒙特卡洛采样算法实现对突变运动的目标跟踪。与文献[90]不同，作者显式地构建一个包含态密度项的试验分布替代原始的滤波分布，并使用 Metropolis-Hastings[95]算法从这个试验分布中采样。从本质上来看，这种采样方式同样也源自于 Multicanonical 采样方法，因而具备有效避免陷入局部模式的重要特性；另一方面，它同时属于动态加权重要性采样算法[97,98]的范畴，样本的权重在采样过程中是动态更新的，这有别于序贯重要性采样算法的粒子的权重估计是一次性获得的。动态加权的优点在于，通过使用一个在线的随机逼近过程(Stochastic Approximation Process, SAP)迭代地估计态密度项，使样本权重估计的精度可以在采样过程中得到持续地提升。此外，作者在蒙特卡洛采样过程中通过引入一个在线预测模型，提出一个新的自适应马氏链蒙特卡洛采样算法以进一步提高整体采样效率。作者的实验表明，该方法能够有效地跟踪多种运动突变复杂场景中的运动目标。

从采样方式来看，Wang-Landau 算法与 SAMC 采样算法都属于自适应马氏链蒙特卡洛(Adaptive MCMC)采样算法[100-103]的范畴。Adaptive MCMC 算法的提出是为了克服传统 MCMC 算法普遍存在的缓慢的收敛速度，提高采样的混合性。需要注意的是，无论是 Metropolis-Hastings 算法[95,111]还是 Gibbs 采样算法[96]，当应用于多模分布采样问题时，往往样本的接受率都很低，尤其在高维复杂多模情况下，采样往往停留在某个样本点或长时间陷入局部模式无法实现模式跳跃(Modes Hopping)。事实上，这与提议分布的选择有很大的关联。一种理想的情况是，尽可能选择或定义一种具有自校正机制的提议分布，使它尽可能地靠近实际的目标分布。这种自校正机制，也就是自适应性，可以利用马氏链上已产生的样本信息来辅助更新提议分布的参数。这就是 Adaptive MCMC 采样算法的基本思想。值得注意的是，无限制地使用样本信息更新提议分布将会破坏转移核的马尔科夫性，从而破坏马氏链的平稳分布。一个可行的方案是，在初始采样阶段实行提议分布的自适应更新，使估计的提议分布较为接近目标分布，而在后续的采样阶段则采用标准的 MCMC 算法从而确保马氏链收敛。Gelfand 和 Sahu[104]通过模拟多条并行的马氏链，使用采样重要性重采样(SIR)算法[74]选择具有高接受率的转移核而抑制那些低接受率的转移核。这种方法的基本出发点是使用近似的马氏链边缘密度作为转移密度。Browne 和 Draper[105]讨论了通过监测接受率的变化调整转移核的部分(或单个)成分，使得采样过程中确保一定的接受概率，提高采样的混合性。由

于 MCMC 算法的有效性很大程度上依赖于马尔科夫转移核的设计，然而如何有效地选择提议分布以及精确地调节其参数仍然是个极具挑战性的课题，这是因为在实际应用中往往无法预知目标分布的形状，尤其是在高维的状态空间中。近年来，在统计学等相关领域自适应 MCMC 算法无论是在理论研究还是在实际应用方面都获得了广泛的关注，已经成为 MCMC 研究领域里一个新的研究热点。其中最重要的理论进展包括：Gilks 等人[99]以及 Brockwell 和 Kadane[106]提出使用再生(Regenerative)马尔科夫链，它在每个再生时刻更新提议分布的参数；Harrio 等人[101]提出一种自适应的 Metropolis 采样算法，它利用马氏链上的所有历史样本信息更新提议分布的协方差矩阵。此外，Andrieu 和 Robert[107]，Atchadé 和 Rosenthal[100]，以及 Roberts 和 Rosenthal[102,103]等人为自适应 MCMC 算法的理论进展和应用做出了重要贡献，包括收敛率和各态历经性(Ergodicity)的理论证明。

总之，在计算机视觉研究领域，国内外基于自适应 MCMC 采样的研究工作还很少。在视觉跟踪研究领域，尽管国内外对运动突变条件下的视觉跟踪的研究已经取得了一些重要成果，但是在这复杂运动条件下实现有效的视觉跟踪仍然是个巨大的挑战，而这又是来源于计算机视觉领域的实际需求而又不得不解决的课题。本研究正是从这一问题的特点和难点入手，基于自适应采样的理论和指导思想探索有效解决这一问题的理论方法和技术。

第 2 章

MCMC和自适应MCMC

2.1 引　　言

随机模拟是利用计算机进行数值计算的一种概率统计方法，也称为统计模拟方法。它通常通过统计试验或随机采样求解数学、物理、工程技术等不同领域问题的近似数值解。随机模拟应用极为广泛，既可以用来解决随机问题，也可以用于求解确定性的科学问题。随机模拟的思想，最早可以追溯到1777年法国数学家Buffon设计实验求解圆周率π。真正意义上使用随机模拟方法解决科学问题是在20世纪40年代。当时它被用于模拟裂变物质的中子链式反应，一项与原子弹研究计划有关的秘密计划。物理学家Von Neumann等人把这种随机模拟方法称为蒙特卡洛(Monte Carlo)方法。

在科学研究中真正系统地运用Monte Carlo方法进行随机试验是在计算机诞生之后。在20世纪50年代，统计物理学家Metropolis等人[111]提出了马氏链蒙特卡洛(MCMC)方法，使随机模拟真正地应用到各个不同的科学研究领域。1970年，Hastings[95]对Metropolis算法进行了推广并提出了至今仍得到广泛应用的Metropolis-Hastings算法。随机模拟技术与计算机科学的发展息息相关，20世纪80年代以后随着计算机科学技术的迅速发展极大地推动了Monte Carlo技术的研究和发展，它在组合优化、贝叶斯推断、非参数统计推断、生物统计学等方面得到了广泛的应用。1984年，S. German和D. German[96]提出了Gibbs采样算法，这是一种基于条件分布的迭代采样方法，它已成为统计计算的一种标准工具。

随着随机模拟与试验设计在各个学科领域的广泛应用，许多领域的复杂科

学问题对蒙特卡洛技术的要求也不断提高。传统的 MCMC 方法对高维复杂分布进行采样时往往遇到很大困难，需要更为有效的采样算法与采样策略来提高随机采样效率。近年来，自适应马氏链蒙特卡洛（Adaptive MCMC）采样方法[99~107]的研究与进展，有效地克服了传统 MCMC 算法的缺点与不足，为 MCMC 方法在统计计算中的进一步应用奠定了坚实的基础。

本章主要介绍 MCMC 和 Adaptive MCMC 的基本概念、理论和算法实现，以及它们在计算机视觉中的应用。

2.2 Monte Carlo 采样

2.2.1 Monte Carlo 积分

在统计推断领域，很多感兴趣的统计量通常可以表示为某随机变量 x 的函数 $h(x)$ 的期望 $E[h(x)]$。令 $f(x)$ 表示变量的概率密度，μ 表示 $h(x)$ 关于 f 的期望。如果能够从 f 抽样一组 i.i.d 样本 $x_1, x_2, \cdots x_n$，那么就可以使用这组样本的均值来近似 μ：

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} h(x_i) \xrightarrow{n \to \infty} \mu \tag{2.1}$$

这里 $\mu = \int h(x) f(x) \mathrm{d}x$。类似地，方差 σ^2 可以通过

$$\widehat{\operatorname{var}}\{\hat{\mu}\} = \frac{1}{n-1}\sum_{i=1}^{n}[h(x_i) - \hat{\mu}]^2 \tag{2.2}$$

来估计。需要指出的是，Monte Carlo 求积法是否收敛，取决于 $|h(x)|$ 的期望是否存在。

在统计计算中，Monte Carlo 求积估计有着广泛的应用。例如，贝叶斯分析中的后验矩估计、后验概率计算、边缘密度计算以及其他一些依赖于随机偏差产生的统计算法等。Monte Carlo 积分收敛速度为 $n^{-(r-1)/r}$，$1 \leqslant r \leqslant 2$，因而 Monte Carlo 收敛速度慢于 $O(n^{-1/2})$。尽管其他一些数值积分方法能够取得更快的收敛速度，但是在高维空间中数值直积分法往往受限于维度问题。Monte Carlo 方法在样本空间中随机采样目标分布，但并非对整个空间进行穷尽探索，因而对样本空间维度的敏感性较低。但是，当维度很高时，Monte Carlo 模拟需要大量的样本来近似实际目标分布，因而也需要很高的计算代价。



2.2.2　Monte Carlo 采样

对于一个随机系统，其中的一些物理量可能不具有确定值，而是满足某个特定的概率分布。Monte Carlo 方法的基本思想就是通过一组服从特定分布的样本点，来估计感兴趣的物理量的统计分布律。物理上的实验粒子方法，就是使用一个粒子集（轨道）来描述一个统计集成（Ensemble）。因而，实现 Monte Carlo 计算的首要工作就是如何从一个已知分布中抽取样本。

我们称想要采样的概率分布 f 为目标分布。如果目标分布来自一个标准参数簇，那么有很多可用于产生样本点的统计软件和工具。通常，这些软件和工具都依赖于标准均匀分布产生随机数。本节主要讨论当目标分布不属于标准参数簇时如何产生实现对随机变量的模拟，包括两种基本的采样方法：拒绝采样（Rejection Sampling）方法[108]和重要性采样（Importance Sampling，IS）[109, 110]方法。

1. 拒绝采样

如果目标分布 f 在仅仅差一个比例常数的情况下是已知的，那么使用拒绝采样从 f 中准确采样是可行的。这种采样方法依赖于一个分布较为简单的辅助分布函数，通过它来修正采样概率。令 q 表示这个辅助概率分布，它通常选择为易于采样和计算的。令 $e(x)=Mq(x)$ 表示一条包络（Envelopcurve），它满足：

$$f(x) \leqslant Mq(x), M > 1 \tag{2.3}$$

基于以上表示，拒绝采样算法可以描述为：

（1）从辅助分布 q 中抽取样本 $y \sim q(\cdot)$；

（2）抽取随机数 $u \sim \mathrm{Unif}(0,1)$；

（3）若 $f(y)/e(y)<u$，则拒绝 y，并返回步骤（1）；

（4）否则接受 y 为目标分布的一个样本，即 $x=y$，然后返回步骤（1），直至抽取到所需数量的样本。

值得注意的是，拒绝采样算法抽取的样本点属于服从目标分布 f 的 i. i. d 样本，这个采样过程并没有引入近似估计，因而采样分布是精确的。此外，对于目标分布 f 在仅差一个比例常数 k 是已知的情况（例如在 Bayes 分析中 f 通常是一个后验分布，它等比例于先验密度和观测似然的乘积），这时如果我们能比较容易地计算 $g(x)=f(x)/k$，那么在这种情况下也可以使用拒绝采样算法，并且采样分布也是精确的。这是因为：按照拒绝采样的思想，如果我们能找到一条包络，满足对所有 x 有 $g(x) \leqslant e(x)$，那么当 $u<g(y)/e(y)$ 时采样被接受，否则采样被拒绝。在

这种情形下采样比例仍然是正确的，因为当 f 被 g 取代时，未知常数 k 在式(2.3)中被抵消了。

与一元分布的情况相似，通过构造一个多元包络，多元目标分布也可以使用拒绝采样算法，算法思想在实现上是相同的。当然，要构造一条合适的包络来覆盖多元目标分布并不容易，这需要对目标分布有足够的了解，确保包络对目标分布的覆盖。一般来说，拒绝采样算法中对包络的要求包括：易于构造和采样，使它能较好地覆盖目标分布；较低的样本拒绝率。

理论上，拒绝采样算法可以精确地对一般的目标分布进行采样，但是，它存在一个主要缺陷，那就是在实际应用中很难在整个样本空间确定合适的包络常数 M。一方面，过大的 M 值尽管可以确保对目标分布的覆盖，但此时的样本接受率

$$\Pr(\text{y is accepted}) = \Pr\left(u < \frac{f(x)}{Mq(x)}\right) = \frac{1}{M} \tag{2.4}$$

变得非常低，这就使拒绝采样方法在高维情形下很难取得满意的成效。

2. 重要性采样

在 Monte Carlo 求积法中，如果使用直接采样，那么被积函数估值所使用的样本点是以均匀的概率抽取的，完全不考虑函数本身的属性特征。假定被积函数在样本点处临近区域有一个陡峭的峰(Peak)，目标分布的尾部(Tail)对于均值估计的贡献很小。直观上看，在对均值估计起重要贡献的区域应该抽取更多的样本。这就要求寻求一个与被积函数的分布情况较为相似的辅助分布函数，转而从中抽取样本，从而提高估值精度。这就是重要性采样的基本思想，它基于这样的一个思路：即 $h(x)$ 关于概率分布 f 的期望可以表示为

$$\mu = \int h(x) f(x) \mathrm{d}x = \int h(x) \frac{f(x)}{g(x)} g(x) \mathrm{d}x \tag{2.5}$$

或

$$\mu = \frac{\int h(x) f(x) \mathrm{d}x}{\int f(x) \mathrm{d}x} = \frac{\int h(x) \frac{f(x)}{g(x)} g(x) \mathrm{d}x}{\int \frac{f(x)}{g(x)} g(x) \mathrm{d}x} \tag{2.6}$$

其中辅助函数 $g(x)$ 称为重要性采样函数或重要性函数。

假定按照重要性采样方法从重要性函数 g 中抽取 n 个 i. i. d 样本点 x_1, x_2, …, x_n，那么 $h(x)$ 关于概率分布 f 的期望可以按下式进行估计：

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} h(x_i) w(x_i) \tag{2.7}$$



其中，$w(x_i)=f(x_i)/g(x_i)$，是样本 x_i 的权重系数。当目标分布 f 是仅差一个比例常数下已知时，应该考虑使用标准化权重 $w^*(x_i)=w(x_i)/\sum_{i=1}^{n}w(x_i)$，其中比例常数在归一化时被抵消，正如前面所述，这种归一化加权方式对 Bayes 分析非常重要。

如何选择重要性函数 g 是个较为复杂的问题。事实上，一个基本的准则是 $f(x)/g(x)$有界同时 g 的尾部比 f 的重(Heavy)。一旦这一准则未能得到满足，将可能导致样本间重要性权重的巨大差别。特别地，如果一个从 g 抽取的样本在 f 下的密度值远高于在 g 下的值，那么它的权重将变得异常大，这将导致很大的估计方差。Liu 等人[74]提出有效样本量(Effective Sample Size，ESS)来度量重要性函数的效率。当使用未归一化权重时，ESS 可以定义为

$$N_{\mathrm{ESS}}=\frac{n}{1+\widehat{\mathrm{var}}\{w^*(x)\}} \tag{2.8}$$

直观上看，有效样本量 ESS 是反映重要性函数 g 与目标分布 f 差异的一个非正式度量，它也可以解释为重要性采样算法中所用的 n 个加权样本相当于 N_{ESS}个真正从 f 中抽取并用于 Monte Carlo 估计的无权重(Unweighted)i. i. d 样本。

2.3　MCMC 采样

在 2.2.2 小节已经介绍了，当直接对目标分布进行采样比较困难时，我们可以使用拒绝采样或重要性采样方法来抽取一个样本集。这类方法抽取的样本主要用于估计某一函数的期望。在这一节将介绍使用马氏链蒙特卡洛(MCMC)方法生成近似服从目标分布的样本。MCMC 方法与前面一节介绍的随机抽样方法的主要区别在于其迭代模拟的特点，而且能够适应各种困难的采样问题。作为一个通用的采样框架，MCMC 对于问题维度的增加不会降低其收敛速度。

MCMC 方法的精髓在于构造一条适当的链，使它收敛于某一个极限平稳分布。令样本序列 $x^{(1)},x^{(2)},\cdots,x^{(n)}$表示一条长度为 n 的马氏链，其状态空间可以是连续的或离散的。以离散状态空间为例，如果 $x^{(i)}$满足：

$$p(x^{(i)}\mid x^{(i-1)},x^{(i-2)},\cdots,x^{(1)})=T(x^{(i)}\mid x^{(i-1)}) \tag{2.9}$$

那么称随机过程 $x^{(i)}$为一条马氏链。如果马氏链在状态空间中的演变(Evolution)仅仅依赖于当前链的状态并且使用固定的转移矩阵，即 $T\triangleq T(x^{(i)}|x^{(i-1)})$，那么称



这种链为齐次马氏链(Hmogeneous Markov Chain)。图 2-1 所示为离散状态空间中一条齐次马氏链状态转移的示意,它基于一个固定不变的转移矩阵。当这条马氏链是非周期(Aperiodicity)不可约(Irreducibility)时,$x^{(i)}$ 收敛到该链的极限平稳分布(非周期是指,对于链上的任何状态,都能以大于零的概率转移到另外一个状态;不可约则表示链不应该陷入循环模式)。MCMC 方法的基本采样策略就是构造一条非周期不可约的马氏链,使其平稳分布对应于给定的目标分布 f。当这条链足够长时,该马氏链获得的样本 $x^{(i)}$ 近似于目标 f 的边际分布。构造这样一条链,一个充分但非必要的条件是可逆(Reversibility)或细致平衡(Detail Balance)条件,即

$$p(x^{(i)})T(x^{(i-1)} \mid x^{(i)}) = p(x^{(x-1)})T(x^{(i)} \mid x^{(i-1)}) \tag{2.10}$$

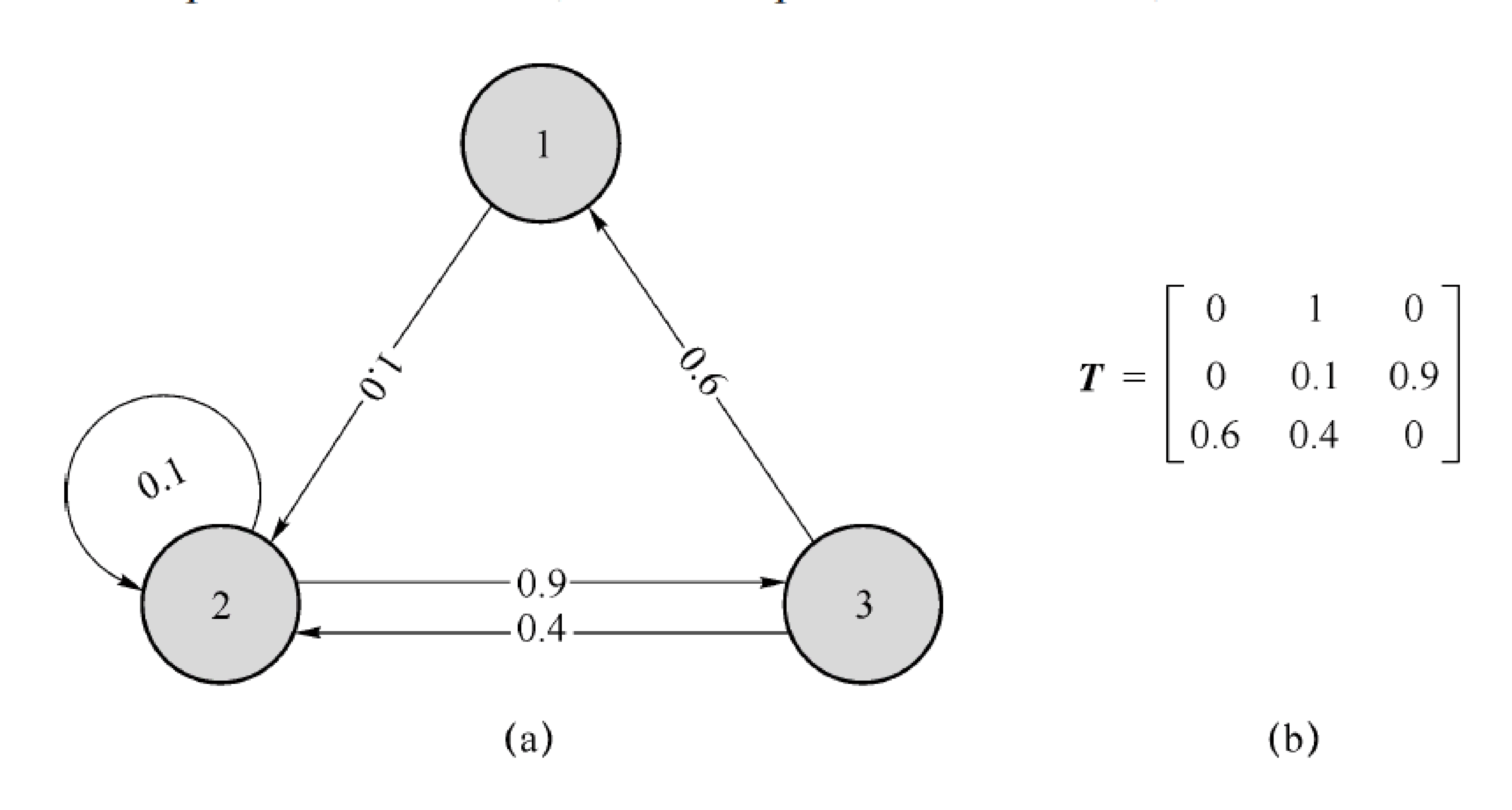


图 2-1　齐次马氏链的状态转移示意及其转移矩阵 T

(a)齐次马氏链的状态转移示意;(b)转移矩阵 T

当状态空间非离散而是连续的时候,转移矩阵 $\boldsymbol{T}$ 就成为了对应的转移核(Kernel) K,转移概率 $p(x)$对应于特征函数:

$$p(x^{(i)}) = \int p(x^{(i-1)})K(x^{(i)} \mid x^{(i-1)})\mathrm{d}x^{(i-1)} \tag{2.11}$$

式中,核 K 是条件密度。

MCMC 采样理论及其应用是统计计算领域的一个研究热点,至今已有大量适用于不同领域和特定科学问题的采样算法。这一节重点介绍一些基本的 MCMC 算法,包括:Metropolis-Hastings 算法[111,95],Gibbs 算法[96],以及它们的变体。这些 MCMC 算法是研究和实现更为高级的 MCMC 算法(例如,自适应 MCMC 方法)的基础,在众多学科领域有着广泛的应用。本节介绍的 MCMC 算法都是基于齐次马氏链构建,在 2.4 节介绍自适应 MCMC 时将讨论非齐次(Non-Hmogeneous)的情况。



2.3.1　Metropolis-Hastings 采样

1. 基本的 Metropolis-Hastings 链

Metropolis-Hastings 算法是一种流行的构造马氏链的通用方法。对于给定的目标分布 f，该算法起始于某个初始样本点 $x^{(0)}$，迭代地从一个提议分布(Proposal Distribution)抽取一系列样本点构建一条马氏链。假定在当前时刻 i，马氏链处于状态 $x^{(i)}$，Metropolis-Hastings 算法的一次迭代产生一个新的样本 $x^{(i+1)}$，算法描述如下：

(1) 从提议分布 $q(\cdot;x^{(i)})$ 中抽取一个候选样本 y；

(2) 计算接受比(Acceptance Rate)$\alpha(y;x^{(i)})$，其中

$$\alpha(y;x^{(i)}) = \min\left\{1,\frac{f(y)q(x^{(i)};y)}{f(x^{(i)})q(y;x^{(i)})}\right\} \tag{2.12}$$

(3) 抽取随机数 $u \sim Unif(0,1)$。若 $u<\alpha$，则接受 y 为新样本，即 $x^{(i+1)}=y$；否则拒绝新样本，即 $x^{(i+1)}=x^{(i)}$。

直观地看，Metropolis-Hastings 算法构造的马氏链完全符合马尔科夫性。而提议分布的设计与选择决定了该链是否是非周期不可约的，如果这个条件得到验证，那么该算法产生的样本将收敛于极限平稳分布 f。

如果提议分布是对称的，即 $q(y;x)=q(x;y)$，那么 Metropolis-Hastings 算法退化为标准的 Metropolis 算法[111]，此时有：

$$\alpha(y;x^{(i)}) = \min\left\{1,\frac{f(y)}{f(x^{(i)})}\right\} \tag{2.13}$$

在通常情况下，对于特定的问题，一个具有特定领域先验知识的提议分布有助于提高 Metropolis-Hastings 算法的采样性能。一个富含信息的(Informative)提议分布能够通过适量的样本覆盖平均目标分布支撑。此外，这样的提议分布的有效性在于它不会导致候选样本被过度频繁地接受或拒绝。多数 MCMC 采样算法的经验结果表明：如果一个提议分布产生的候选样本过于分散(比如在全局样本空间内随机跳跃)，那么由此产生的候选样本就会被频繁的拒绝，从而导致马氏链需要很长时间的迭代才能充分地获得目标分布的支撑空间的粗略模式(Rough Mode)；与此相反，如果提议分布过于集中(比如使用较小方差的高斯分布)，那么尽管接受率较高，但马氏链在多次迭代后无法跳跃到样本空间中别的局部模式，从而不能有效地对其他区域进行充分地采样。尤其在高维情形下，采样效率对提议分布的依赖程度将变得更为明显。

对于不同的提议分布设计方案，Metropolis-Hastings 算法的实现对应着不同的采样器，本节后面部分将介绍两种常用的 Metropolis-Hastings 算法变体，包括：独立 Metropolis-Hastings 算法，随机游走 Metropolis（Random Walk Metropolis，RWM）算法以及多步试验 Metropolis-Hastings 算法[112]。

2. 独立 Metropolis-Hastings 算法

在 Metropolis-Hastings 算法中，如果提议分布使用一个固定的函数使得它产生的候选样本 y 不依赖于链上的当前样本 x，即 $q(y;x)=q(y)$，那么由此构建的马氏链称为独立链。这种情形下，样本接受比为

$$\alpha(y;x^{(i)}) = \min\left\{1,\frac{f(y)q(x^{(i)})}{f(x^{(i)})q(y)}\right\} \tag{2.14}$$

很明显，如果 $f(x)>0$ 且 $q(x)>0$，那么这条独立链就是非周期不可约的，它将收敛于极限平稳分布 f。独立 Metropolis-Hastings 算法是一种易于实现的简单采样方法，因而它在统计学中有着广泛的应用，尤其是在 Bayes 分析中，常常通过构建一条独立链来估计后验矩和其他一些统计量，此时 Bayes 先验分布通常被选择为提议分布。此外，通过简单的变换，可以建立起独立 Metropolis-Hastings 算法与重要性采样算法在概念上的关联。事实上，式(2.14)可以改写为

$$\alpha(y;x^{(i)}) = \min\left\{1,\frac{f(y)/q(y)}{f(x^{(i)})/q(x^{(i)})}\right\} = \min\left\{1,\frac{w(y)}{w(x^{(i)})}\right\} \tag{2.15}$$

也就是说，此时的样本接受比相当于样本的重要性权重比。这表明，在独立链的构建中，应该选择那些能够覆盖目标分布大部分支撑区域的提议分布，否则独立链的接受率将变成很低，并长时间停留在某个样本点。

3. 随机游走 Metropolis 算法

随机游走算法是最为常用的 MCMC 算法之一，是 Metropolis-Hastings 算法的一个特例，由此算法构建的马氏链也称为随机游走链。与独立链不同，随机游走链的样本生成依赖于当前的样本状态。假定在当前时刻 i，马氏链的当前状态为 $x^{(i)}$，候选样本的生成按照以下方式：

$$y = x^{(i)} + \varepsilon,\varepsilon \sim g(\varepsilon) \tag{2.16}$$

其中，$g(\varepsilon)$为一随机密度。对于随机游走链，提议分布具有以下特征：

$$q(y;x^{(i)}) = h(y-x^{(i)}) \tag{2.17}$$

通常，函数 h 可以选择为高斯扰动，即 $h:\mathcal{N}(0,\sigma^2)$，其中 σ 为高斯扰动方差。易于验证由此算法构建的随机游走链是非周期不可约的。

4. 多步试验 Metropolis-Hastings 算法

Metropolis-Hastings 算法的主要困难在于选择合适的提议分布，一个不恰当



的提议分布可能导致马氏链的收敛速度很慢，长时间易于陷入局部模式而无法从中跳出。一个朴素而直接的方案就是扩大提议分布的采样范围(Range)，使提议的候选样本尽可能命中(Hit)目标分布的支撑区域。但是，这种方式由于盲目的在庞大样本空间中进行探索往往造成极低的样本接受率，尤其是在高维的状态空间中。针对这种矛盾，Liu 等人提出一种通过扩大提议分布的采样范围，但同时不影响采样混合性的多试验 Metropolis-Hastings 算法[112]。

假定在当前时刻 i，马氏链的状态为 $x^{(i)}$，多步试验 Metropolis-Hastings 算法的一次迭代产生一个新样本 $x^{(i+1)}$，算法描述如下：

(1) 从提议分布 $q(\cdot;x^{(i)})$中抽取 m 个 i. i. d 候选样本：$y_1,y_2,\cdots,y_m$；

(2) 从 m 个候选样本中，以概率 $w(x^{(i)},y_j)$

$$w(a,b) = f(b)q(a;b)\lambda(a,b) \tag{2.18}$$

抽取候选样本 y_j，其中 $\lambda(a,b)$为一非负的加权函数；

(3) 给定候选样本 y_j，使用提议分布 $q(\cdot;y_j)$抽取 $m-1$ 个 i. i. d 样本：y'_1，$y'_2,\cdots,y'_{m-1}$，同时令 $y'_m=y_j$；

(4) 计算采样接受比 $\alpha(y_j;x^{(i)})$：

$$\alpha(y_j;x^{(i)}) = \min\left\{1,\frac{\sum_{i=1}^{m} w(x^{(i)},y_i)}{\sum_{i=1}^{m} w(y_j,y'_i)}\right\} \tag{2.19}$$

(5) 以概率 $\alpha(y_j;x^{(i)})$接受 y_j 为新样本，即令 $x^{(i+1)}=y_j$；否则拒绝 y_j，并令 $x^{(i+1)}=x^{(i)}$。

多步试验 Metropolis-Hastings 算法通过提议分布生成一定量的候选样本，并基于这些候选样本产生一个新样本。其基本思想是通过一定量的候选样本探究目标分布 f 在当前样本 $x^{(i)}$附近的分布情况，从而构建一条马氏链，使链上的样本具有较小的相关性(Correlation)，以此来提高采样的混合性。相比于前面所述的其他几种 MCMC 算法，该采样算法在理论上似乎更完美一些。但需要注意的是，多个候选样本的权值计算意味着更大的计算代价，尤其是基于 Bayes 分析的视觉计算应用中(例如，图像分割和视觉跟踪等)，权值的计算往往涉及代价很高的似然计算。

2.3.2　Gibbs 采样

前面所述的 Metropolis-Hastings 算法并未对样本空间的维度进行专门的讨

论。对于多维目标分布，Gibbs 采样[96]是最为常用的一种 MCMC 算法，它通过目标分布的边缘分布序贯抽取样本来构造一条收敛于目标分布的马氏链。

对于 p 维随机变量 $x=(x_1,x_2,\cdots,x_p)^{\mathrm{T}}$，记 $x_{-i}=(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_p)^{\mathrm{T}}$，假定目标分布 f 的条件密度 $f(x_i|x_{-i})$，$i=1,2,\cdots,p$ 有显式的表达式且易于采样，那么简单的(Plain)Gibbs 采样过程可描述如下：

(1) 初始化 $x^{(0)}=(x_1^{(0)},x_2^{(0)},\cdots,x_p^{(0)})^{\mathrm{T}}$；

(2) For $i=0$ to $N-1$

- 抽取 $x_1^{(i+1)}\sim f(x_1|x_2^{(i)},x_3^{(i)},\cdots,x_p^{(i)})$；
- 抽取 $x_2^{(i+1)}\sim f(x_2|x_1^{(i+1)},x_3^{(i)},\cdots,x_p^{(i)})$；

⋮

- 抽取 $x_j^{(i+1)}\sim f(x_j|x_1^{(i+1)},\cdots,x_{j-1}^{(i+1)},x_{j+1}^{(i)}\cdots,x_p^{(i)})$；

⋮

- 抽取 $x_p^{(i+1)}\sim f(x_p|x_1^{(i+1)},x_2^{(i+1)},\cdots,x_{p-1}^{(i+1)})$；

End For.

在 Gibbs 采样中，一个很重要的特点是序贯地采用最新值。S. German 和 D. German[96]证明了 $x_j^{(i)}$ 的极限边缘分布等于目标分布 f 的第 j 维边缘分布，并且由 Gibbs 采样生成的样本可以用于估计随机变量 x 的函数的期望。事实上，Gibbs 采样算法是 Metropolis-Hastings 算法的一个特例。在 Metropolis－ Hastings 算法中，对于 x 的每一维 $j=1,2,\cdots,p$，采用以下的提议分布：

$$q(x;x^{(i)})=\begin{cases} f(x_j \mid x_{-j}^{(i)}) & x_{-j}=x_{-j}^{(i)} \\ 0 & \text{其他} \end{cases} \tag{2.20}$$

如果易于验证此时的样本接受比等于 1，那么提议产生的候选样本总是被接受。

对于目标分布 f，当不是所有的条件密度都有显示的表达式或不易于抽样时，混合 Gibbs 采样和 Metropolis-Hastings 采样的方法能够有效地构建 Gibbs 链。这类方法包括 Metropolis-within-Gibbs[113]和格点 Gibbs 采样[114]等。

2.3.3 MCMC 算法分析

在前面的章节里我们介绍了马氏链的基本构建方法和相关算法，本节主要讨论 MCMC 算法的收敛、混合性及其诊断等问题。

给定目标分布，MCMC 算法的可靠性很大程度上依赖于马氏链的样本序列是否收敛于极限平稳分布以及收敛速度。前面介绍的 Metropolis-Hastings 算法与 Gibbs 算法都能够收敛于其极限平稳分布。但是，在实际应用中还必须对所构建



的马氏链有充分的了解，以便知道该链何时能够输出足够代表目标分布的样本，从而对感兴趣的统计量进行可靠的估计。在对复杂目标分布进行采样时，MCMC 算法往往收敛缓慢，需要长时间的模拟和迭代才能接近目标分布，特别是当样本空间维度很高的情形下。因而，如何实施 MCMC 算法以确保马氏链的收敛以及良好的混合性就显得尤为重要。在 MCMC 采样算法的实际运行中，我们首先需要考虑的是马氏链是否已近似地收敛到其平稳目标分布。当前，已有一些现成的分析方法可用于诊断马氏链的收敛性[127,128]。但值得注意的是，对不同马氏链的收敛性诊断，很难找到一种通用的分析方法和工具。此外，MCMC 算法还需考虑链的混合性，它通常可以通过两个方面来反映，即马氏链上的样本对目标分布支撑区域的"覆盖"程度，以及马氏链上近似独立的样本点间的间隔距离。

首先，我们讨论提议分布的设计与选择对马氏链收敛和混合性的影响。在前面的介绍中我们已经提到了提议分布的延展度对 MCMC 算法的混合性有重要的影响，同时也应该认识到提议分布的设计与选择也在一定程度上依赖于特定的 MCMC 算法。对于基本的 Gibbs 采样算法，如果样本状态的各维分量之间独立性越高其影响就越大。降低 Gibbs 链相关性提高混合性的途径包括：重新参数化(Reparameterization)[75]和使用辅助变量等方法。对于 Metropolis-Hastings 链，通常的思路是尽可能选择那些与目标分布较接近的提议分布，从而可以获得较高的样本接受率。即便如此，提议分布的尾部分布情况也同样对链的混合性有着重要影响。总的来说，提议分布的选择与设计最好是使其在分布上比目标分布更为分散一些，同时使得目标分布与提议分布之比在样本空间内有界[113]。在接受率对链的混合性的影响方面，Roberts 等人指出当目标分布为二维（不含）以上的高斯时，随机游走 Metropolis 算法的最佳接受率为 0.234，而当目标分布维度不大于 2 时最佳的接受率大约为 45%[119]。而对于多步试验 Metropolis-Hastings 算法，Liu 等人[112]指出链的接受率在 40%～50%之间时有助于算法取得良好的混合性。值得注意的是，对于接受率的设定在一定程度上也依赖于特定的目标分布，当目标分布为复杂多模时，若能获知其有关先验知识无疑对提议分布的选择以及接受率的设定都有重要的指导意义。

其次，我们讨论马氏链的个数对 MCMC 算法收敛速度的影响。在对复杂多模分布进行随机模拟时，往往很难判定链是否已经近似地收敛于平稳分布，这是由于 MCMC 采样可能陷入局部模式。在这种情形下，考虑构建多条马氏链是个较为理想的选择，通过综合分析多条链的收敛情况可能有助于估计模拟产生的样本与目标分布的接近程度[129]。事实上，学术界在使用单一长链还是多条短链存在着一些

争议。使用多条短链的观点认为，通过多条链共同对目标分布进行探索有助于提高对多模分布模拟的成功率，进而降低链的初始状态对收敛的影响。而使用单一长链的观点则认为，使用多条短链在诊断链收敛不佳时确实比较有用，但这些短链往往也是不稳定的。此外，在给定迭代次数(计算资源)的情况下，将计算资源集中用于一条长链可能会获得较好的收敛，而把迭代分给多条短链则可能难以获得满意的成效。即便存在争议，一个不容忽视的事实是，由多条不同初始状态构建的马氏链通常能够为我们提供一些有关未知目标分布的有用信息，这些信息有助于后续构建长链，以便更有效地探索目标分布。

最后，我们讨论马氏链的丢弃长度(Burn-in Perieod)以及运行长度。鉴于任何 MCMC 算法，只有当链的运行长度为无穷大时，才真正收敛于平稳分布，因而一定长度的链都只是近似地服从边缘分布。此外，由于有限的模拟次数，马氏链对初始状态通常也比较敏感。考虑这些因素，在实际应用中 MCMC 算法通常考虑丢弃链上最初的一定长度的样本。这些丢弃样本的数目称为丢弃长度。Gelman-Rubin 方法[130,131]是最为常用的用于估计丢弃长度与链运行长度的一种方法，它通过运行多条链并利用链内方差与链间方差的差异来确定丢弃长度和链长。假定 MCMC 算法使用 m 个长度都为 n 的链对目标分布进行模拟，记丢弃长度为 n_b，第 j 条链上的第 i 个样本为 $x_j^{(i)}$，同时令

$$\bar{x}_j = \frac{1}{n-n_b}\sum_{i=n_b}^{n-1} x_j^{(i)}, \quad \bar{x} = \frac{1}{m}\sum_{j=1}^{m}\bar{x}_j \tag{2.21}$$

链间方差为

$$S_b = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{x}_j - \bar{x})^2 \tag{2.22}$$

链内平均方差为

$$\bar{S} = \frac{1}{m}\sum_{j=1}^{m} S_j \tag{2.23}$$

式中，S_j 表示第 j 条链的链内方差，即

$$S_j = \frac{1}{n-n_b}\sum_{i=n_b}^{n-1}(\bar{x}_j^{(i)} - \bar{x}_j)^2 \tag{2.24}$$

Gelman-Rubin 统计量 $\hat{R}$ 定义为

$$\sqrt{\hat{R}} = \sqrt{\frac{\frac{n-n_b-1}{n-n_b}\bar{S} + \frac{1}{n-n_b}S_b}{\bar{S}}} \tag{2.25}$$



从式(2.25)可以看到，如果链间方差较大，则 $\hat{R}>1$。随着模拟的进行，$\lim\limits_{n\to\infty}\hat{R}=1$。Gelman 等人[131]建议使用阈值$\sqrt{\hat{R}}=1.1$ 作为马氏链近似收敛的判定准则。若不满足该条件，可通过增加链长(或丢弃长度)继续模拟。此外，评估链的收敛与混合性还可通过一些定性分析工具和方法，包括样本路径图和自相关性图(Autocorrelation)等。

2.4　自适应 MCMC 采样

通过 2.3 节的介绍我们知道，传统的 MCMC 算法在采样过程中通常使用固定不变的提议分布。无论是 Metropolis-Hastings 算法还是 Gibbs 采样算法，以及它们的变体，在对复杂多模的目标分布进行采样时往往样本的接受率都较低，尤其在高维的情形下，采样可能长时间停留在某个点无法跳跃到别的模式。事实上，这与提议分布的设计与选择有很大的关系。

自适应 MCMC 采样的目的正是为了克服传统 MCMC 算法普遍存在的缓慢的收敛速度，提高采样的混合性。直观上看，应该尽可能选择或设计一种具有自校正机制的提议分布，使它尽可能地接近目标分布的形状。这种自校正机制，也就是自适应性(Adaptibility)，通常可通过马氏链上的历史样本信息来实现。利用马氏链上的已有样本信息来更新和估计提议分布是自适应 MCMC 算法的基本手段。值得注意的是，无限制地使用样本信息更新提议分布将会破坏转移核的马尔科夫性，从而破坏马氏链的平稳分布。由于 MCMC 算法的有效性很大程度上依赖于马尔科夫转移核的设计，然而如何有效地选择提议分布以及精确地调节其参数仍然是个极具挑战性的课题，这是因为在实际应用中往往无法预测目标分布的形状，尤其是在高维的状态空间中。近年来，在统计学等相关领域自适应 MCMC 算法无论是在理论研究还是在实际应用方面都获得了广泛的关注，已经成为 MCMC 研究领域里一个新的研究热点。按照自适应采样的方式，自适应 MCMC 可以归结为三大类：基于再生的自适应[99,106]，提议分布自适应[100~103,117]，以及目标分布自适应[91,92,122,93]。基于再生的自适应 MCMC 方法，主要由 Gilks 等人[99]、Brockwell 和 Kadane[106]提出并发展的。再生马氏链使在再生点之后生成的样本独立于再生点之前的历史样本，这种提议分布的更新能够保证马氏链的收敛，但是该类算法在实际应用中采样效率并不理想，而且如何确定再生点也是一项困难的任务。本节主要介绍更为常用、相对易于实施的提议分布和目标分布自适应算法。

2.4.1 提议分布自适应

这类 MCMC 算法试图在采样过程中在线学习一个高效的提议分布，也就是说在每次迭代时都利用已有的样本更新提议分布。Haario 等人[101]提出一种自适应的 Metropolis 采样算法，它利用马氏链上的所有历史样本信息更新提议分布的协方差矩阵。作者证明了在一定条件下，随着采样过程的进行，在线学习到的提议分布将收敛于一个“最优”提议分布。随后，Andrieu 和 Robert[107]阐明了该自适应算法实际上属于 Ronbbins-Monro 类型[115]的一种随机逼近（Stochastic Approximation）算法。Atchadé 和 Rosenthal[100]在随机逼近理论的框架下，证明了更为一般的自适应 MCMC 算法的收敛性。随后，Roberts 和 Rosenthal[102、103]以及 Yang[116]进一步发展了自适应 MCMC 算法理论，并给出了更为简单的条件确保自适应 MCMC 算法的各态历经性。

记 $f(\cdot)$为 σ—场 $\mathcal{F}$ 中 d 维样本空间 $\chi\in\mathbb{R}^d$ 上的一个不变的目标分布，$\{K_\theta\}_{\theta\in\Theta}$为样本空间$\mathcal{X}$上以 f 为平稳分布的马氏链转移核的一个集合，即$(fK_\theta)(\cdot)=f(\cdot)$，其中，$\Theta$ 为一参数空间。假定 K_θ 是非周期不可约的，那么

$$\lim_{n\to\infty}\|K_\theta^n(\cdot;x)-f(\cdot)\|=0,\forall x\in\mathcal{X} \tag{2.26}$$

其中，$\|u(\cdot)-v(\cdot)\|=\sup_{A\in\mathcal{F}}\|u(A)-v(A)\|$ 代表全变分距离。这意味着，如果转移核不变，即 θ 保持固定不变，由此转移核生成的样本将收敛于目标分布 f。对于 Adaptive MCMC 算法，允许转移核在采样过程中动态的改变，在提高采样的混合性的同时仍能保证收敛于目标分布。记 $\theta=\Gamma^{(n)}$ 为 Θ 型的一个随机变量，它确定了在采样过程中第 n 次迭代所使用的转移核 K_θ，假定在第 n 次迭代时链上的样本状态为 $x^{(n)}$，那么

$$K_\theta(A;x)=p(x^{(n+1)}=A\mid x^{(n)}=x,\Gamma^{(n)}=\theta,A\in\mathcal{F},S^{(n)}),A\in\mathcal{F} \tag{2.27}$$

其中，$S^{(n)}=\sigma\{x^{(1)},x^{(2)},\cdots,x^{(n)};\Gamma^{(1)},\Gamma^{(2)},\cdots,\Gamma^{(n)}\}$，是由集合$\{x^{(i)},\Gamma^{(i)}\}_{i\leqslant n}$生成的一个筛（Filtration）。给定初始状态 $x^{(1)}$ 和初始核 $\Gamma^{(1)}$，记第 n 次迭代时样本状态为 $x^{(n)}$ 的条件密度为

$$A^{(n)}(B;(x^{(1)},\Gamma^{(1)}))=p(x^{(n)}\in B\mid x^{(1)},\Gamma^{(1)}),B\in\mathcal{F} \tag{2.28}$$

并记 $x^{(n)}$ 与目标分布 f 的全变分距离为

$$\|A^{(n)}(\cdot;(x^{(1)},\Gamma^{(1)}))-f(\cdot)\|=\sup_{B\in\mathcal{F}}\|A^{(n)}(B;(x^{(1)},\Gamma^{(1)}))-f(B)\| \tag{2.29}$$



对于这一类型的自适应 MCMC 链，Roberts 和 Rosenthal[102,103]给出了确保链的各态历经性的充分条件，即有界收敛条件(Bounded Convergence Condition)和渐减自适应条件(Diminishing Adaptation Condition)。这里，有界收敛条件指的是：给定初始样本状态 $x^{(1)}$ 和初始核 $\Gamma^{(1)}$，对任意的 $\varepsilon>0$，随机过程$\{t_\varepsilon(x^{(n)},\Gamma^{(n)})\}$依概率有界，其中

$$t_\varepsilon(x,\theta)=\inf\{n\geqslant 1:\|K_\theta^{(n)}(\cdot;x)-f(\cdot)\|\leqslant\varepsilon\} \tag{2.30}$$

表示起始于(x,θ)的马氏链的收敛时间。

渐减自适应条件指的是：$\lim\limits_{n\to\infty}D_n=0$，其中

$$D_n=\sup_{x\in\mathcal{X}}\|K_{\Gamma^{(n+1)}}(\cdot;x)-K_{\Gamma^{(n)}}(\cdot;x)\| \tag{2.31}$$

渐减自适应条件要求随着采样的进行，对转移核参数的校正必须逐渐减小，或者以逐渐减小的概率进行自适应采样[118]。Roberts 和 Rosenthal[102]给出了关于基于提议分布自适应的 MCMC 链收敛的基本定理。

定理 2.1　对于一个基于自适应提议分布的 MCMC 算法，若满足有界收敛条件和渐减自适应条件，则由此构建的马氏链将收敛于平稳分布 f。

本节主要介绍两个重要且常用的基于提议分布自适应的 MCMC 采样算法：Adaptive Metropolis(AM)算法[101]和 Adaptive Independent Metropolis-Hastings (AIMH)算法[117]。

1. AM 算法

在前一节中介绍的随机游走 Metropolis 采样算法中，如果提议分布 $q(y;x)$设定为均值为 x、协方差为Σ的多元高斯分布$\mathcal{N}(y;x,\Sigma)$，现在考查如何对该随机游走链进行自适应构建。对于提议分布的协方差矩阵，过大或过小都将可能导致链的高相关性，因而如何设计和选择协方差矩阵显得尤为重要。Gelman 等人[119]指出随机游走链的“最优”协方差矩阵为

$$\Sigma_{\mathrm{opt}}=(2.38^2/d_x)\Sigma_{\mathrm{f}} \tag{2.32}$$

式中，d_x 为样本 x 的维度；Σ_f为目标分布的协方差矩阵。该协方差通常情况下是未知的，Andrieu 和 Moulines[120]指出，为计算Σ_f，可通过最小化 Kullback-Leibler 距离

$$\int_{\mathcal{X}} f(x)\log\frac{f(x)}{q(x)}\mathrm{d}x=\mathbb{E}_f\left\{\log\frac{f(x)}{\mathcal{N}(x;\mu,\Sigma)}\right\} \tag{2.33}$$

Haario 等人[101]则提出在采样过程中在线估计这个协方差矩阵，并由此协方差矩阵迭代采样的 AM 算法。给定链的初始样本状态 $x^{(1)}$ 和初始提议分布参数 $\mu^{(1)}$，$\Sigma^{(1)}$，AM 算法描述如下：

(1) 初始化样本状态 $x^{(1)}$ 和转移核参数 $\theta^{(1)}=(\mu^{(1)},\Sigma^{(1)})$;

(2) For $i=1$ to $N-1$

– 使用 Metropolis-Hastings 转移核 $K_{\theta^{(i)}}(\cdot;x^{(i)})$ 抽取一个新样本 $x^{(i+1)}$,其中 $\theta^{(i)}=(\mu^{(i)},\Sigma^{(i)})$;

– 更新转移核参数

$$\mu^{(k+1)}=\mu^{(k)}+\lambda_{k+1}(x^{(k+1)}-\mu^{(k)}), \tag{2.34}$$

$$\Sigma^{(k+1)}=\Sigma^{(k)}+\lambda_{k+1}\{(x^{(k+1)}-\mu^{(k)})(x^{(k+1)}-\mu^{(k)})^T-\Sigma^{(k)}\} \tag{2.35}$$

式中,λ_k 为学习率(Learning Rate)。

End For.

尽管 AM 算法试图在线地估计和学习目标分布的协方差,在高维情形下算法所估计的协方差在某些方向上不是过大就是过小。这使得 AM 采样算法的接受率不是过高就是过低,从而影响对目标分布协方差的准确估计。究其原因,主要是因为采样算法只是对样本空间的有限范围进行了探索。事实上,对于高维情形下的 AM 采样算法,当目标分布的成分之间近似不相关(Uncorrelated)时,Roberts 和 Rosenthal[121]指出最优的样本接受率应该为 0.234,过高或过低的接受率都不利于提高采样的混合性。

2. Adaptive Independent Metropolis-Hastings 算法

AIMH 算法与前面介绍的独立 Metropolis-Hastings 算法较相似,区别在于前者考虑使用自适应的提议分布,而后者使用固定不变的转移核。同时,AIHM 算法也不同于前述的 AM 算法,它不需要通过渐减自适应方式来确保马氏链的各态历经性。Holden 等人[117]提出使用马氏链上的部分样本来更新提议分布,同时避免渐减自适应条件的要求。

记 $q^{(n)}(\cdot;X_{n-1})$ 为 AIMH 算法在第 n 次迭代时使用的提议分布,其中 X_{n-1} 表示提议分布所使用的马氏链上的历史样本集。鉴于 AIMH 采样算法试图构建一条独立马氏链,因而样本集合 X_{n-1} 不应包含链上的当前样本点 $x^{(n)}$,即 $x^{(n)}\notin X_{n-1}$。对于历史样本集 X_n 的构建与更新,AIMH 算法采用的方案是:假定候选样本 $y\sim q^{(n)}(\cdot;X_{n-1})$,如果该样本被独立链接受为新样本,那么将 $x^{(n)}$ 添加到历史样本集 X_{n-1} 中构成新的集合 X_n;否则将候选样本 y 添加到历史样本集 X_{n-1} 中构成新的集合 X_n。AIMH 算法描述如下:

(1) 初始化样本状态 $x^{(1)}$,转移核参数 $\theta^{(1)}$;

(2) 初始化历史样本集 $X_1=\varnothing$;

(3) For $i=1$ to $N-1$



– 使用提议分布 $q^{(n)}(\cdot \mid X_i)$抽取一个候选样本 y；

– 计算 Metropolis-Hastings 接受比 $\alpha(y, x^{(i)}, X_i)$

$$\alpha(y, x^{(i)}, X_i) = \min\left\{1, \frac{f(y)q^{(i)}(x^{(i)} \mid X_i)}{f(x^{(i)}q^{(i)}(y \mid X_i)}\right\} \tag{2.36}$$

– 以概率 $\alpha(y, x^{(i)}, X_i)$接受 y 为新样本，令 $x^{(i+1)} = y$；否则拒绝 y，并令 $x^{(i+1)} = x^{(i)}$。

– 若 y 被接受，则令 $X_{i+1} = X_i \cup \{x^{(i)}\}$；否则令 $X_{i+1} = X_i \cup \{y\}$。

End For.

定理 2.2　自适应独立 Metropolis-Hastings 算法构建的马氏链收敛于目标分布 f，亦即，如果 $p_i(x^{(i)} \mid X_{i-1}) = f(x^{(i)})$，那么有 $p_i(x^{(i+1)} \mid X_i) = f(x^{(i+1)})$。

Holden 等人[117]证明了 AIMH 算法构建的独立链以目标分布 f 为其极限平稳分布。记 $\tilde{p}_n(X_n)$为马氏链在第 n 次迭代时历史样本集的联合分布，w 为第 $n+1$ 次迭代时添加到历史样本集 X_n 中的样本。事实上，假定 $p_n(x^{(n)} \mid X_{n-1} = f(x^{(n)})$，则

$$\begin{aligned} p_{n+1}(x^{(n+1)} \mid X_n)\tilde{p}_n(X_n) &= \tilde{p}_{n-1}(X_{n-1})\{f(w)q^{(n)}(x^{(n+1)} \mid X_{n-1})\alpha(x^{(n+1)}, w, X_{n-1}) \\ &\quad + f(x^{(n+1)})q^{(n)}(w \mid X_{n-1})(1 - \alpha(w, x^{(n+1)}, X_{n-1}))\} \\ &= \tilde{p}_{n-1}(X_{n-1})\{f(x^{(n+1)})q^{(n)}(w \mid X_{n-1}) + f(w)q^{(n)}(x^{(n+1)} \mid X_{n-1})\alpha(x^{(n+1)}, w, X_{n-1}) \\ &\quad - f(x^{(n+1)})q^{(n)}(w \mid X_{n-1})\alpha(w, x^{(n+1)}, X_{n-1})\} \\ &= f(x^{(n+1)})q^{(n)}(w \mid X_{n-1})\tilde{p}_{n-1}(X_{n-1}) \end{aligned}$$

因而 $p_n(x^{(n+1)} \mid X_n) = f(x^{(n+1)})$。定理 2.2 得证。

2.4.2　目标分布自适应

在自适应 MCMC 算法簇中，有一类算法在构建马氏链时并不考虑自适应调整和学习转移核，而是自适应的更新目标分布自身。也就是说，这类算法试图在采用过程中迭代地更新其试验分布（Trial Distribution），使马氏链最终收敛于目标分布。我们称这类算法为基于目标分布自适应的 MCMC 算法。这类自适应算法的典型代表包括：Multicanonical 采样算法[92]，Wang-Landau 算法[91]，广义 Wang-Landau 算法[122]和随机逼近 MCMC（SAMC）算法[93]等。本节主要介绍前两个算法，将在第 3 章中详细介绍随机逼近 MCMC 采样算法及其在视觉跟踪中的应用。

事实上，Wang-Landau 算法，广义 Wang-Landau 算法可以看作是 Multicanonical 采样算法的变体，区别在于如何构建试验分布以及所采用的学习策略。值得注意的是，此类算法构建试验分布的初衷是为了克服传统 MCMC 算法（包括简单的 Gibbs 采样和 Metropolis-Hastings 算法）在对复杂多模的目标分布采样时往往易于陷入局部模式的缺陷，这类算法常用于解决统计物理学中一些物理量的模拟和计算[92,93]。类似于前面介绍的基于提议分布自适应的 MCMC 算法，基于目标分布自适应的 MCMC 算法也利用马氏链上的历史样本信息，来更新和指导试验分布的学习。

假定目标分布 f 可以表示为仅差一个比例常数 c 的已知分布，即

$$f(x) = cp(x), x \in \mathcal{X} \tag{2.37}$$

其中 $\mathcal{X}$ 为样本空间。记 $U(x) = -\log p(x)$，它在物理学中表示分布 $p(x)$ 对应的能量函数。当目标分布的能量地形（Energy Landscape）崎岖不平（rugged）时，传统的 MCMC 采样算法往往表现出混合性差，长时间停留在某个局部模式难以靠近或到达全局最优模式。1991 年，Berg 和 Neuhaus[92] 提出了用于一阶相变过程中态密度（Density of States, DoS）估计的一种 MCMC 采样算法。

1. Multicanonical 采样算法

在物理学中，态密度也称为光谱密度（Spectral Density），它用于描述一个系统中每一能量阶对应的能量子区域（Subregion）中可用状态的数目。某个特定能量阶的 DoS 较高，意味着对应子区域有较多的可用状态。态密度的估计有助于对其他物理量的估计和计算，它在物理中也用于解释许多量子力学现象[91]。

假设使用 Multicanonical 采样算法从一个正则系综（Ensemble）抽取样本。系统的每一个位型（Configuration）x 都被赋予一个权重 w：

$$w(u) \sim \frac{1}{g(u)} = \mathrm{e}^{-S(u)} \tag{2.38}$$

式中，$u = U(x)$ 表示位型 x 的能量阶，$S(u)$ 称为微正则系综熵。如果使用权重函数式（2.38）进行随机模拟，将使能量的分布 $p(u) \sim g(u)w(u) \sim 1$。这也意味着这种随机采样将在能量空间进行随机游走，并能够访问到能量空间的各个子区域。同时，由此产生的样本在能量空间中形成一个平直的直方图，因而 Multicanonical 算法也称为平直直方图（Flat Histogram）蒙特卡洛算法[93]。

现在考虑离散状态空间的情形。在此情况下位型 x 的能量可以在有限个能量阶中取值，即 $U(x) \in \{u_1, u_2, \cdots, u_m\}$。假定在第 n 次迭代后抽取的样本序列为 $\{x^{(i)}\}_{i=1}^{n}$，记 $n_T(i) = \#\{x^{(j)}: U(x^{(j)} = u_i)\}$，则



$$\lim_{n\to\infty}\frac{n_T(i)}{n}=\frac{1}{Z(T)}g(u_i)\mathrm{e}^{-u_i/T},i=1,2,\cdots,m \tag{2.39}$$

式中，$Z(T)$为配分函数，T 为系统的温度。因而态密度可以按下式进行估计：

$$g(u_i)=\frac{n_T(i)\mathrm{e}^{u_i/T}}{\sum_{j=1}^{m}n_T(j)\mathrm{e}^{u_j/T}},i=1,2,\cdots,m \tag{2.40}$$

Multicanonical 采样算法分多个阶段进行，在完成一个阶段采样后，更新一次态密度项，也就是权重项。整个采样过程的算法描述如下：

(1) 初始化态密度 $g_0(u_i)$，$i=1,2,\cdots,m$；$s=0$；

(2) For $s=1$ to N

- 使用 Metropolis-Hastings 算法从权重函数 $w^{(s)}(x)$中抽取 N_s 个样本，其中

$$w^{(s)}(x)=\frac{1}{g_s(U(x))} \tag{2.41}$$

- 重新估计态密度

$$g_{s+1}(u_i)=c(\phi_s(i)+\alpha_i)g_s(u_i),i=1,2,\cdots,m, \tag{2.42}$$

式中，$\phi_s(i)$表示第 i 个能量阶的相对采样频率；α_i 为平滑项；常数 c 确保所估计的态密度值在有效的范围内。

End For.

Multicanonical 采样算法的有效性依赖于其态密度的学习机制。事实上，Multicanonical 采样算法对态密度的估计通常只能取得有限的统计精度，当采样到达一定阶段之后，继续增加采样阶段或采样次数将无助于态密度估计的进一步改进[93]。

2. Wang-Landau 采样算法

Wang-Landau 算法是 Multicanonical 采样算法的一个变体，两者的主要区别在于估计态密度的学习算法不同。在随机逼近的理论框架下，两者都可以看作 Robbins-Monro 算法[115]的一个特例，Wang-Landau 算法采用的学习策略较之 Multicanonical 算法采样更为有效，因而对态密度的估计能够取得更好的统计精度。Wang-Landau 算法的最重要的一个特征是能够有效地避免在采样过程中陷入局部能量极小点，这个特征对于具有复杂能量地形特点的物理系统来说至关重要。近年来，它在统计物理学中得到了极为广泛的应用，已成为用于态密度等物理量估计的最重要、最常用的统计计算方法之一。

假设使用 Wang-Landau 算法从一个正则系综抽取样本。为系综的每一个位型 x 都赋予一个权重 w：

$$w(u) \sim \frac{1}{g(u)} \tag{2.43}$$

式中，$u=U(x)$，表示位型 x 的能量阶；$g(u)$ 为态密度。在离散状态空间中，位型 x 的能量可以在有限个能量阶中取值，即 $U(x)\in\{u_1,u_2,\cdots,u_m\}$。Multicanonical 算法类似，Wang-Landau 算法也包括多个采样阶段。在每个采样阶段中，迭代地使用 Metropolis-Hastings 算法抽取样本序列直到在能量空间中获得一个平直直方图为止，然后采样进入下一个阶段。如果每个能量阶的采样频率不低于能量阶的平均采样频率的 80%，则这个直方图被认为是平直的。一旦获得一个平直的直方图，算法就对态密度学习的增益因子进行更新，通常是选择渐减的一个增益因子序列。算法重复进行这样的采样过程，直到增益因子接近于 1。Wang-Landau 算法描述如下：

(1) 初始化 ξ,ε；态密度 $g(u_i)=1, i=1,2,\cdots,m$；

(2) Repeat

- Repeat
  - 使用 Metropolis-Hastings 算法从试验函数 $f_w(x)$ 中抽取一个样本 x，其中

$$f_w(x) = \frac{1}{g(U(x))} \tag{2.44}$$

  - 更新态密度

$$g(u_i) = g(u_i)\xi^{I(U(x)=u_i)}, i = 1,2,\cdots,m \tag{2.45}$$

- Until 获得一个平直的能量直方图；
- $\xi \leftarrow \sqrt{\xi}$

Until $\xi \leqslant \varepsilon$

尽管 Wang-Landau 采样算法在对许多物理系统的随机模拟与计算中取得了很大的成功，但是至今还没有充足的理论证实它的收敛性，而且在一些统计计算中已经发现它只能取得有限的统计精度，也就是说当采样达到一定阶段之后，持续的迭代采样不能进一步改进统计量的估计精度[123]。Liang 等人[93]研究了 Wang-Landau 算法的缺陷与不足，并指出其主要原因在于增益因子的设计与选择策略。紧接着，Liang[122]扩展并提出了一个广义的 Wang-Landau 算法。这种扩展主要包括：将离散空间推广到连续空间；引入能量空间划分的概念；以及更为有效的增益因子更新策略。随后，Liang 等人[93]在随机逼近理论框架下[115,124]进一步扩展和改进了广义 Wang-Landau 算法，确保算法在理论上的收敛性，使算法的应用范围延



展到物理学以外的一些统计计算领域，例如模型选择、高后验区间估计、随机优化等。

2.5 MCMC 采样方法在计算机视觉中的应用

本节主要介绍 MCMC 采样方法在计算机视觉中的应用。特别地，探讨了它们在目标轮廓近似表示中的应用。在随机优化的框架下，提出了一种基于 Monte Carlo 全局优化的用于轮廓表示的多边形近似(Polygonal Approximation)算法[125]和一种基于自适应 MCMC 采样的多边形近似算法[126]。实验表明，提出的算法与同类的主流算法相比具有更好的近似精度代价比。

2.5.1 目标轮廓的多边形近似

多边形近似是图像目标轮廓近似表示的一种重要方法，它在计算机视觉、模式识别、信号处理等领域有着重要的应用。在多边形近似算法中，一条封闭的平面曲线被分割成首尾连接的一系列线段，这些连接的线段构成一个多边形近似地表示这条封闭曲线。多边形近似方案的优点在于通过尽可能地保留轮廓曲线的关键特征而忽略那些冗余的曲线点，从而节省存储空间并减低在应用中的处理计算代价。同时，它可以通过增减特征点来平衡近似方案的性价比，即近似精度与计算(或存储)代价的平衡，因而具有较好的灵活性。

对于一条封闭的平面数字曲线，多边形近似问题可以归为两类，即最小近似错误(min−ε)问题和最小线段数目(min−#)问题。前者的任务是在给定线段数目的前提下寻找一个具有最小近似错误的多边形，后者的任务是在给定最大允许近似错误的情况下寻找一个具有最少线段数的多边形。对于多边形近似问题，现有的处理方法大致可以归为两大类：基于局部搜索的算法[132,133]和基于全局搜索的算法[125,126,135-139]。

局部搜索的方法主要是利用曲线在数据点(Data Point)的局部信息，比如曲率等信息，迭代地寻求局部最佳近似方案。局部搜索方法的典型代表包括分裂-合并(Split-and-Merge)算法[132]、控制点检测(Dominant-point Detection)算法[133]等。分裂-合并算法交互地对一个随机的初始近似多边形迭代地进行分裂与合并操作，直到获得算法所能取得的最佳方案为止。在每一个分裂阶段，多边形的一条线段被其对应弧上的某个数据点分裂为两条连接的线段。执行分裂操作的数据点的选

择依赖于一个策略：在所有数据点中选择一个距离其对应线段最远的一个作为分裂点。而在每一个合并阶段，多边形中某两条互相连接的线段被合并成一条线段，并在多边形中删除该顶点以及与其相连的两条线段。删除点的选择准则是：在多边形的所有顶点中选择一个离合并后线段距离最近的顶点。一般来说，分裂一合并算法对近似多边形的初始方案比较敏感，因而它获得的近似方案可能与实际最佳方案相去甚远。控制点检测算法的基本思想是通过对曲线的局部属性（主要指的是曲率）分析，获得一系列控制点，以此来构建一条近似多边形。这种算法的近似性能与曲线的局部信息提取有很大的关系，局部信息的精度对近似方案有关键性的影响。

不同于局部搜索方法，基于全局的近似方法试图通过全局搜索获得全局最佳多边形近似方案。基于全局的搜索算法主要是基于元启发（Meta-Heuristics）的方式，包括：基于遗传的近似算法[135-137]、基于粒子群优化的近似算法[139]以及基于蚁群优化的算法[138]等。实践证明，基于元启发的全局近似算法在多边形近似中能够取得较好的近似精度代价比。一般来说，在近似精度方面，基于全局搜索的方法比基于局部搜索的算法表现更佳，并且对初始方案不敏感；但是全局搜索往往需要更大的计算代价。因而，一个有效的近似算法需要在近似精度与计算代价之间求得一种平衡，取得最佳的近似精度代价比。事实上，多边形近似问题可以通过动态规划等一些方法求得精确的最佳方案[134]，然而这种方法通常需要付出极高的计算代价，而这在实际应用中通常是不可行的，尤其对于有着大数据点的数字曲线近似问题。

本节将 MCMC 采样方法引入到多边形近似中，提出一种基于 Monte Carlo 优化的多边形近似算法[125]和一种基于自适应 MCMC 采样的多边形近似算法[126]。在介绍提出的算法之前，我们先给出多边形近似的问题定义。

令 $C:(p_1,p_2,\cdots,p_n)$ 表示一条具有 n 个有序数据点的封闭平面曲线，其中 $p_1=p_n$。记这条数字曲线的多边形近似方案为 $P:x=(x_1,x_2,\cdots,x_n)$，任一候选方案 x 为一个二值向量（Binary Vector），其中

$$x_i=\begin{cases}1 & p_i\in P\\ 0 & \text{其他}\end{cases} \tag{2.46}$$

于是，两类多边形近似问题可以定义为：

min－ε：给定一条平面数字曲线 $C:(p_1,p_2,\cdots,p_n)$，以及固定的多边形顶点数 m，寻找一个多边形 $P:x$，使得在所有可能的近似方案里该多边形与曲线 C 之间有最小的近似错误；



min－#：给定一条平面数字曲线 C：$(p_1,p_2,\cdots,p_n)$，以及近似错误上限 ε，寻找一个多边形 P：x，使得在所有可能的近似方案里该多边形的顶点数最少。

多边形近似问题中，通常用积分平方错误(Integral Square Error, ISE)来度量多边形与曲线间的近似错误。令 $\widehat{p_ip_j}=\{p_i,p_{i+1},\cdots,p_j\}$ 表示起点和终点分别为 p_i 和 p_j 的顺时针方向上曲线的一段弧，$\overline{p_ip_j}$ 表示连接数据点 p_i 和 p_j 的一条线段，$d(p_k,\overline{p_ip_j})$ 表示数据点 p_k 与线段 $\overline{p_ip_j}$ 之间的垂直距离。于是，近似错误 ISE 可以表示为

$$\varepsilon(\widehat{p_ip_j},\overline{p_ip_j}) = \sum_{p_k \in \widehat{p_ip_j}} d^2(p_k,\overline{p_ip_j}) \tag{2.47}$$

常用的多边形近似方案度量准则包括压缩比和轮廓应用。

(1) 压缩比(Compression Ratio, CR)：

$$\mathrm{CR}=\frac{n}{n_d} \tag{2.48}$$

式中，n 为曲线的数据点个数；n_d 为近似多边形的顶点个数。使用较少顶点的多边形近似能获得高的压缩比，但是其近似精度可能较低。

(2) 轮廓效用(Figure of Merit, FOM)：

Sarkar[140] 提出将压缩比与近似错误两种度量结合起来评价近似方案，即

$$\mathrm{FOM}=\frac{\mathrm{CR}}{\mathrm{ISE}} \tag{2.49}$$

在实际应用中，FOM 和 ISE 是较常用的度量标准。对于 min－ε 问题，由于各种方案都使用固定的多边形顶点数，因而度量比较简单，只需要比较不同方案之间的 ISE 即可；而对于 min－# 问题，可通过比较 FOM 值评价不同的方案，FOM 值越大其性价比越好。

2.5.2　基于 MC 优化的多边形近似算法

1. 方法介绍

在离散优化的框架下，两类多边形近似问题可以形式化地描述为：

(1) min－ε：给定一条数字曲线 C：$(p_1,p_2,\cdots,p_n$，以及多边形顶点数 m，寻找满足下式的方案 x：

$$x = \mathrm{argmin}_x \mathrm{ISE}(x,C),s.t.\sum_{i=1}^{n} x_i = m \text{ and } x_i \in \{0,1\} \tag{2.50}$$

(2) min－#：给定一条数字曲线 C：$(p_1,p_2,\cdots,p_n)$，以及近似错误上限 ε，寻找满足下式的方案 x：

$$x = \text{argmin}_x \sum_{i=1}^{n} x_i, s.t. \text{ISE}(x, C) \leqslant \varepsilon \text{ and } x_i \in \{0,1\} \tag{2.51}$$

两类多边形近似问题都属于带约束的离散优化问题。本节将基于 MC 采样的全局优化和基于分裂-合并过程的局部优化相结合,提出一种混合的离散优化算法来处理两类多边形近似问题。在提出的算法中,通过使用分裂-合并过程去除两类优化问题的约束条件,从而把约束优化转化为无约束的离散优化问题。令 $f(x)$表示离散优化的目标代价函数,其中 x 代表满足约束条件的可行解(方案),则两类优化问题可以统一表示为

$$x = \text{argmin}_x f(x), x \in \mathcal{X} \tag{2.52}$$

式中,$\mathcal{X}$表示可行解空间。

我们考虑采用随机优化方法解决式(2.52)的离散优化问题。特别地,我们采用基于 Basin-Hopping 采样[141]的 MC 优化方法。当应用于全局优化问题时,MC 方法的有效性高度依赖于它是否能够在采样中避免陷入局部模式,并能够在低能量区域进行有效的采样。在实际应用中,目标分布的能量地形往往是崎岖不平的,其中有很多高能量屏障(Barrier),这使许多优化和随机优化方法难以翻越这些能量屏障达到或靠近全局最近模式。Basin-Hopping 采样通过在每次 MC 采样步骤中嵌入一个局部优化过程,来消除能量地形中的高能量屏障,使最终的采样在一个更为"平坦"的能量地形上进行,从而避免采样陷入局部模式,进而提高采样效率。从随机优化的角度来看,这种采样方式使得对低能量区域的搜索变得更为容易,增加了靠近或达到全局最低能量点的机会。Basin-Hopping 采样方法在统计物理和固体物理学中有着广泛的应用,我们首次将这种采样方法引入到多边形近似问题中,并提出一种结合分裂-合并操作的混合优化算法来解决两类多边形近似问题。

在优化的初始阶段,随机选择一个初始的多边形近似方案。从初始点 x 开始,执行一次基于分裂-合并的局部优化过程,状态从 x 移动到局部能量极小点 $x_{\min}$;然后通过一个提议分布,状态跳跃到一个新的方案点 x',然后再执行一次局部优化过程后到达一个新的局部极小点 $x'_{\min}$。最后,算法通过计算 Metropolis 接受率 α:

$$\alpha = \min\left\{1, \frac{q(x; x')}{q(x'; x)} \mathrm{e}^{-\Delta E_{\min}/K_B T}\right\} \tag{2.53}$$

其中,$\Delta E_{\min} = f(x'_{\min}) - f(x_{\min})$;$K_B$ 为波尔兹曼因子;T 为系统的温度。算法的一次采样过程如图 2-2 所示。在采样过程中可以采用模拟退火[142]的方式逐渐降低温度 T,然而在我们的算法中使用的提议分布是全局跳跃的,因而 $K_B T$ 项的影响可以忽略。因而,接受率 α 改写为:



$$\alpha = \min\left\{1, \frac{q(x;x')}{q(x';x)} e^{-\Delta E_{\min}}\right\} \tag{2.54}$$

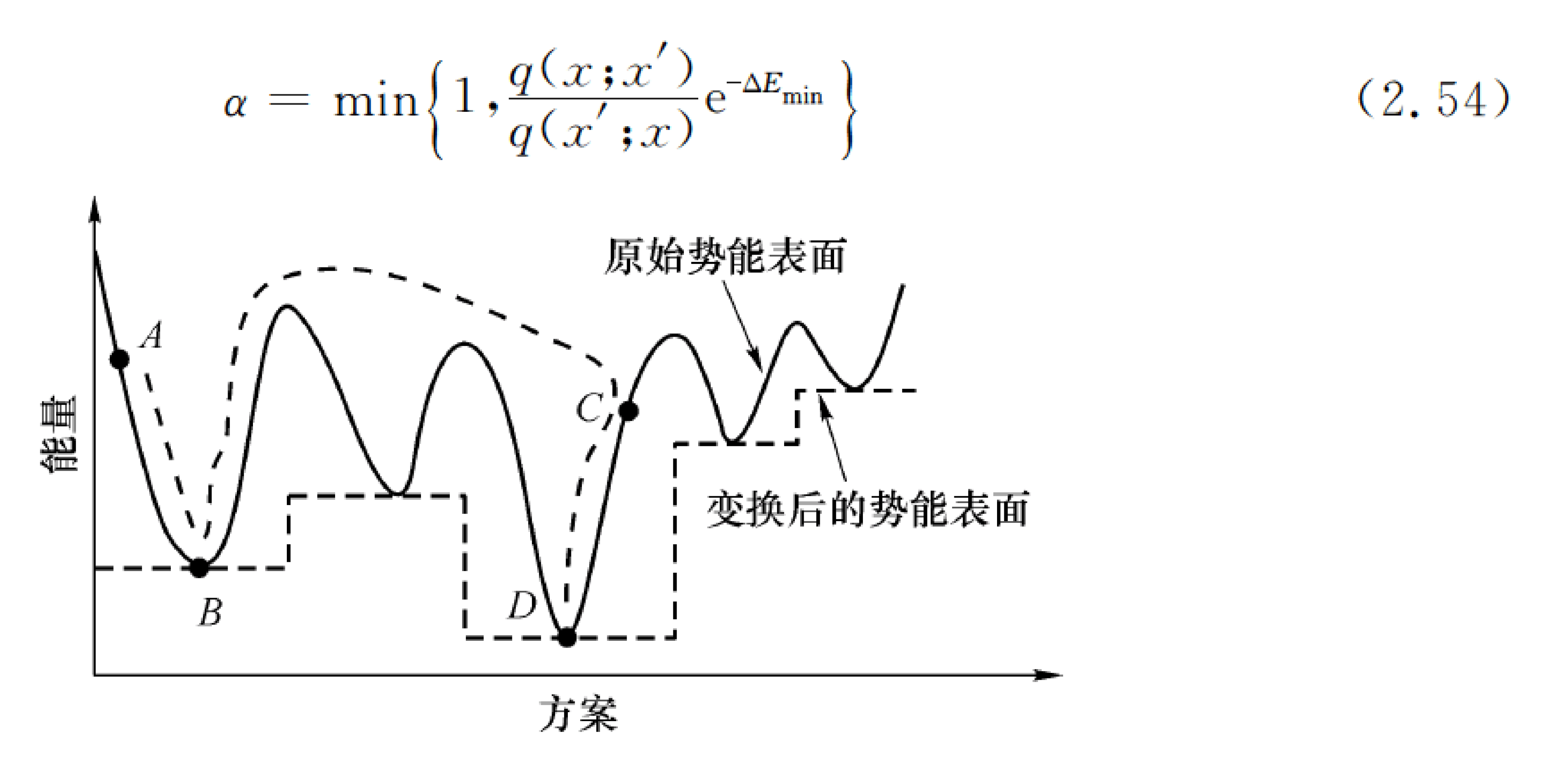


图 2-2　Basin-Hopping 采样方法示意。原始的势能表面经局部优化后变成更为平坦的阶梯形的能量地形分布，使得 MC 采样更易于跨越高能量屏障，避免陷入局部模式。

两类多边形近似问题的 MC 优化算法描述如下：

(1) 在可行解决方案中随机生成一初始方案 x；

(2) 运行一个基于分裂或合并操作的局部优化获得一个极值点 $x_{\min}$；

(3) For $i=1$ to N　//N 表示采样次数

- 基于分裂-合并的提议分布 q 生成一个新的方案 x'；
- 运行一个基于分裂或合并操作的局部优化获得一个极值点 $x'_{\min}$；
- 按照式(2.54)计算接受比 α；
- 以概率 α 接受方案 x'；若方案 x' 被接受，则 $x=x'$；$x_{\min}=x'_{\min}$；否则，保留 x，x'不变；
- $x^{(i)}=x_{\min}$；

End For

(4) $x^*=\operatorname{argmin}_{x^{(i)}} f(x^{(i)})$，$i=1,2,\cdots,N$。

算法中的局部优化过程是基于多边形的分裂或合并操作的。局部优化的任务是把从一个初始方案点移动到该点附近的能量极小点。多边形的分裂操作描述如下：给定一个待分裂的数据点 p，把点 p 所在的弧分割成两段弧，然后把数据点 p 加入到多边形顶点集中，并连接点 p 与相邻两多边形顶点，并把两条边加入到多边形中。分裂操作通过增加多边形的顶点数来降低近似错误 ISE。多边形的合并操作描述：给定一个待合并的多边形顶点 p，把点 p 以及与其相连的两条边从多边形中删除，并把与删除点相邻的两个多边形顶点相连。合并操作通过删除多边形的顶点来增加近似方案的压缩比，但同时加大了近似错误 ISE。多于两类多边形近似问题，它们对应的局部优化过程也不一样，我们分别加以讨论：

对于 min－ε 问题,局部优化描述:给定一个初始方案点,局部优化可通过执行一系列的分裂-合并过程,直到获得局部能量极小点。在这个过程中分裂操作和合并操作是成对出现的,即每一次分裂操作之后执行一次合并操作。在分裂操作中,分裂点的选择原则是:选择一个数据点,使在此处的分裂将导致 ISE 值降低幅度最大。在合并操作中,合并点的选择原则是:选择一个顶点,使在此处的合并操作将导致 ISE 值增加幅度最小。值得注意的是,如果初始方案是在可行解空间中,那么基于分裂-合并操作的局部优化过程不会改变方案的可行性,使方案仍然属于可行解空间,这是由于分裂-合并操作不会改变多边形的顶点数。

对于 min－# 问题,局部优化过程基于一系列合并或分裂操作,但这两种操作并非成对出现。在优化过程中,一系列的合并或分裂操作可能导致候选方案不再属于可行解空间,因而需要考虑方案修复。方案修复同样也是基于合并或修复操作。局部优化算法描述如下:

(1) 给定初始方案 x, 近似错误上界 ε;

(2) If $f(x)>\varepsilon$ then

Repeat

－ 在数据点 x 处执行分裂操作;

Until $f(x)\leqslant\varepsilon$

End If

(3) 令 $x^*=x$, 对方案 x 执行一次合并操作;

(4) 若 $f(x)>\varepsilon$,则返回局部最佳方案 x^*;否则转(3)。

在提出的 MC 优化算法中,使用提议分布进行随机跳跃,从一个方案点跳跃到方案空间中的另一个候选方案点。随机跳跃对 MC 优化算法有重要影响,在我们的算法中,这种随机跳跃是基于分裂-合并操作实现的。有别于局部优化过程中对分裂点和合并点的选择的是确定性的,随机跳跃中对分裂点和合并点的选择基于一个权重系数(概率)。下面介绍在提议分布 q 中分裂点和合并点的选择原则:

(1) 分裂。在选择分裂点时,采样算法倾向于寻找一个“理想”的数据点进行分裂。对每个不在多边形中的数据点 p_i,首先计算其与对应多边形边的垂直距离 d_i,然后以概率 w_i,即

$$w_i=\frac{d_i}{\sum_{i=1}^{n}d_iI(x_i=0)} \tag{2.55}$$

选择数据点 p_i 为分裂点。其中,I 为示性函数。这一分裂点选择方案倾向于距离



多边形边较远的数据点。在实际应用中,也可以考虑选择倾向于有最大 ISE 降幅的分裂点,但是这种方式考虑的还是能量的增减,直接倾向于在低能量区域产生样本点不一定有利于产生全局的随机跳跃,因而我们考虑前一种方案,即使用距离而不是直接使用能量来度量权重。

(2) 合并。在产生一个分裂点之后,紧接着需要产生一个合并点。对每个多边形顶点 p_i,首先计算其与待添加的多边形边(连接与合并点相邻的两个顶点的一条边)的垂直距离 d_i,然后以概率 w_i,即

$$w_i = \frac{d_i}{\sum_{i=1}^{n} d_i I(x_i = 1)} \tag{2.56}$$

选择数据点 p_i 为合并点。其中,I 为示性函数。合并点选择方案倾向于距离多边形边较近的顶点。

2. 实验结果

为评估提出的算法在解决两类多边形问题中的性能,实验中使用了一个公用数据集[134]。数据集包含了用于测试多边形近似方案的几组目标轮廓点集:Figure-8,Chromosome,Semicircle,Leaf 等。

在实验中,我们将使用的算法与当前几种主流的多边形近似算法基于遗传算法的多边形近似算法[135,136]以及基于局部搜索的控制点检测算法[133]进行了比较:在我们的算法中,采样的样本数目设置为 200,而在多边形近似算法[135,136]中,种群(Population)大小和代(Generation)的数目均设置为 100。多边形近似算法和我们的算法都属于基于随机优化算法,它们在每个测试用例上都独立运行 10 次,实验中给出的结果是 10 次运行中的最佳结果。实验所选用的评估准则是 ISE 度量和 FOM 度量,这两种度量在多边形近似问题中得到了广泛的应用,是评估多边形近似算法的最常用的标准度量方式。表 2-1 和表 2-2 给出了不同的多边形近似算法在测试集上的实验结果;图 2-3 给出了 MC 优化算法在测试集上的部分实验结果。

表 2-1　min—# 问题的算法结果比较

Shape	Method	ε	m	ISE	FOM
Figure-8	Wang [136]	6	8	5.90	0.953
	proposed		8	5.90	0.953
	Wang [136]	10	7	7.83	0.821
	proposed		7	7.83	0.821
	Wang [136]	20	6	17.83	0.421
	proposed		6	17.49	0.429

续 表

Shape	Method	ε	m	ISE	FOM
Chromosome	Wang [136]	8	11	7.42	0.735
	proposed		11	7.34	0.743
	Wang [136]	10	10	8.07	0.743
	proposed		10	8.07	0.743
	Wang [136]	20	7	17.68	0.485
	proposed		7	17.68	0.485
	Wang [136]	30	6	23.73	0.421
	proposed		6	23.73	0.421
Semicircle	Wang [136]	20	14	18.16	0.401
	proposed		14	17.39	0.419
	Wang [136]	30	12	27.28	0.312
	proposed		12	26.06	0.326
	Wang [136]	60	10	46.18	0.221
	proposed		10	46.00	0.222

表 2-2　min—ε 问题的算法结果比较

Shape	Method	m	ISE	FOM
Semicircle	Chen [135]	22	7.19	0.645
	Masood [133]		8.61	0.538
	Wang [136]		7.04	0.659
	proposed		7.01	0.661
	Chen [135]	26	4.73	0.829
	Masood [133]		4.91	0.799
	Wang [136]		4.05	0.969
	proposed		4.05	0.969
	Chen [135]	27	3.74	1.010
	Masood [133]		4.40	0.859
	Wang [136]		3.70	1.021
	proposed		3.70	1.021
Leaf	Chen [135]	23	9.87	0.529
	Masood [133]		10.61	0.492
	Wang [136]		9.53	0.547
	proposed		9.46	0.552



续 表

Shape	Method	m	ISE	FOM
Leaf	Chen [135]	32	4.68	0.801
	Masood [133]		4.97	0.755
	Wang [136]		4.45	0.843
	proposed		4.45	0.843

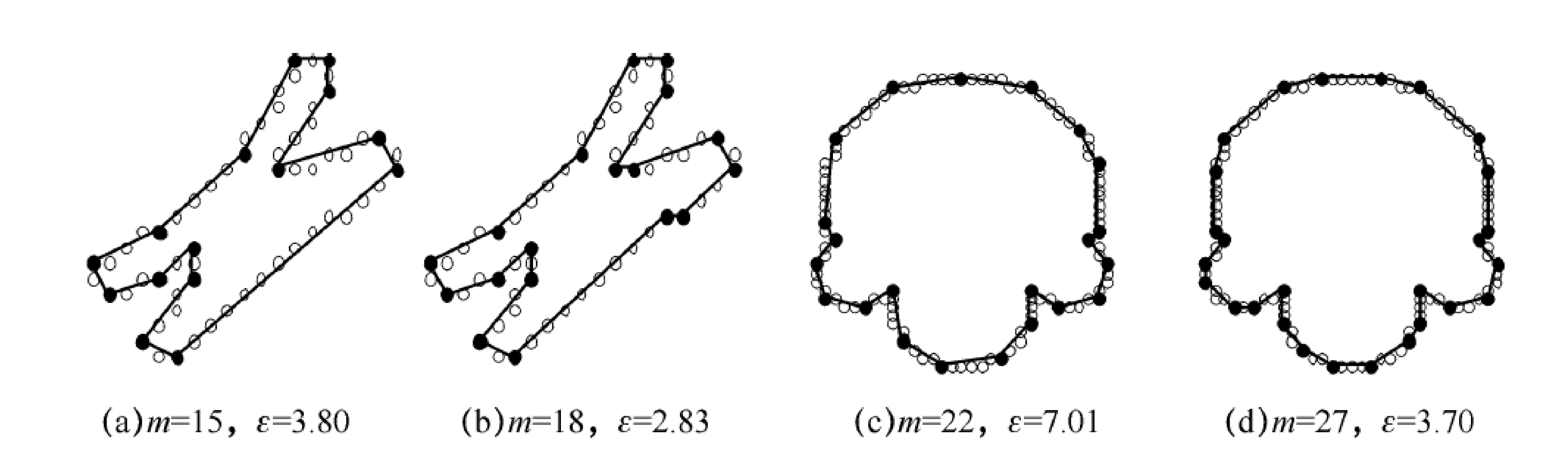


图 2.3　MC 优化算法在数据集 Chromosome(左)和 Semicircle(右)上的多边形近似结果

从中我们可以看到，对于两类多边形近似问题，我们的算法获得了最优或接近最优的方案，且相比其他三种近似算法，我们的算法有更好的近似精度代价比。原因可以解释为：

首先，近似算法[135]和[136]都基于遗传算法，尽管他们具有全局优化算法的特征，能够在方案空间中进行全局搜索，但它们在局部搜索方面缺乏有效性。从能量空间的角度分析，当多边形近似问题的能量分布崎岖不平时，搜索往往易于陷入局部能量极小点。遗传算法通常需要大量的个体(迭代)来实现在方案空间的全局搜索，因而当种群大小和代的数目较小时，很难靠近或获得全局最佳方案。我们的方法尽管使用较低的计算代价，但是通过基于分裂或合并操作的局部优化，MC 采样在变换后的较为平坦的能量空间上进行采样，因而能够有效地避免采样陷入局部模式。此外，通过局部优化过程，提高了整体采样的效率，使得采样算法在较小的迭代次数内访问到不同的局部模式，增加了达到或靠近全局最佳方案的机会。

其次，对于控制点检测算法[133]，尽管它能够通过曲线的局部特征(例如曲率)分析进行近似方案的局部搜索，具有很高的局部搜索性能，然而基于局部搜索的近似算法往往对初始方案点的选择较为敏感，并且它没有有效的机制处理搜索过程中经常出现的局部模式陷入问题(Local-Trap Problem)。因而它不易于获得多边形近似问题的全局最佳方案，尤其是当轮廓曲线的数据点非常庞大的情况下。

我们的算法,本质上属于一类 Basin-Hopping 蒙特卡洛采样方法,它除了具有局部优化能力,还可通过 MC 全局采样指导优化过程,避免优化陷入局部能量极小点。

2.6 小　　结

本章介绍了 MCMC 采样方法的基本概念、理论和算法实现;讨论了 MCMC 算法的实施和性能评估方法。研究了传统的 MCMC 采样算法在高维状态空间中采样时表现出混合性较差且易于陷入局部模式的基本原因。介绍了自适应 MCMC 采样方法的基本概念、理论和算法实现,分析了当前主流的三类 Adaptive MCMC 算法(即再生马氏链、提议分布自适应和目标分布自适应)的采样机制和内在关联。探讨了 MCMC 采样算法在计算机视觉中的应用,提出了一种结合局部优化与全局 MC 采样的目标轮廓近似表示算法。实验表明,基于分裂-合并的局部优化技术与基于采样的全局优化相结合能够有效地解决两类多边形近似问题。与同类的主流算法相比,我们提出的算法具有更好的近似精度代价比。对于更为复杂的多边形近似问题,例如,大规模数据点的轮廓曲线近似等问题,算法还需结合轮廓曲线自身的局部或全局特征信息,设计更为有效地提议分布,从而进一步提高算法的近似性能。

第 3 章

序贯随机逼近蒙特卡洛跟踪方法

3.1 引　　言

作为计算机视觉领域经典的、核心的研究内容之一，视觉跟踪技术融合了图像处理与理解、模式识别、机器学习以及计算机软硬件等众多学科领域的方法与技术。计算机视觉研究具有较高的科学价值和广阔的应用前景，而视觉跟踪是计算机视觉研究领域的基础和关键环节之一，是该领域中更高层次（如行为分析、理解等）研究的基础。

动态场景中的视觉跟踪可看作可视目标在视频的相继帧之间建立一种对应关系，它是计算机视觉研究领域的一项重要研究课题，是视频分析、行为识别和运动分析的基础和关键环节，在众多学科和领域有着广泛的应用。

作为计算机视觉研究中的一个热点研究方向，视觉跟踪在近年来得到了高度的关注和广泛的研究。从方法论的角度，视觉跟踪的研究方法可归结为自底向上和自顶向下的两种思路。自底向上的方法更多依赖于可靠的检测技术，直接从视觉场景中获取目标的相关信息并用于视觉跟踪。自顶向下的方法也称为模型驱动的方法，最典型的例子是线性或非线性的状态空间模型的方法。对应于自底向上和自顶向下的两种研究思路，当前的视觉跟踪算法大抵可以分为两大类：确定性跟踪方法和随机跟踪方法[2]。

确定性跟踪方法一般利用自底向上的底层视觉信息提供可靠的检测结果实现快速的目标跟踪。这些自底向上的视觉信息包括第 1 章介绍的目标特征信息，包括 Haar－ like 特征、颜色直方图、梯度方向直方图（HOG）、边缘方向直方图

(EOG)、尺度不变的特征变换(SIFT)等。尽管基于 Mean Shift 的确定性跟踪方法取得了很大的成功,但是这种方法通常要求在相继的两帧之间目标的位移较小。其主要原因在于 Mean Shift 算法容易受到多模场景中背景噪声等因素干扰而陷入局部最优状态,从而导致跟踪失败。对于运动不确定性的目标跟踪问题,确定性的跟踪方法容易跟踪失败并且很难从错误状态中恢复。尽管结合随机采样或模拟退火过程的 Mean Shift 算法能够在一定程度上提高跟踪的稳健性,但在处理复杂多模分布的问题时仍然难以取得令人满意的跟踪性能。

不同于确定性的跟踪方法,基于随机采样的概率跟踪方法能有效地实现在多模分布场景下的视觉跟踪任务。这类方法通常是在基于贝叶斯跟踪框架下,通过状态预测与观测似然迭代地估计后验分布实现对目标状态的估计。在运动模型与观测度量呈线性,噪声服从高斯分布的情况下,卡尔曼滤波[70]可以获得贝叶斯滤波的精确解析解。对于非高斯条件下的非线性系统,通常是使用 MC 模拟技术来实现迭代的贝叶斯滤波。序贯 MC 算法或粒子滤波[4]是此类跟踪方法的典型代表。一般情况下,使用大量的粒子能够很好地近似复杂的滤波分布,但是高昂的计算代价往往使得跟踪算法不能满足实际的跟踪需要。

本章探讨在贝叶斯滤波框架下的视觉跟踪方法。通过引入一种高效的自适应 MCMC 采样方法到贝叶斯滤波框架,提出一种新的概率跟踪算法,即序贯随机逼近蒙特卡洛(Sequential SAMC,S-SAMC)跟踪算法,它能对复杂多模的滤波分布进行有效地采样。同时,它是一个通用的跟踪框架,易于集成各种有效的观测模型和先验模型以进一步提高跟踪算法的稳健性。

3.2 相关工作

在基于随机采样的视觉跟踪方法中,粒子滤波是最为成功的一个跟踪框架,它具有很好的可扩展性。一般来说,直接从后验(滤波)分布采样是比较困难的。在状态空间中有效的采样策略对于粒子滤波算法来说至关重要。从采样的角度看,粒子滤波算法可以归为三类,即重要性采样[71~74],因子采样[4],以及拒绝采样[75]等。重要性采样的基本出发点在于尽可能从样本空间中靠近后验模式的区域采样。对于贝叶斯滤波问题,重要性采样的关键在于如何选取提议分布。序贯重要性采样(SIS)是最为常用的一种解决贝叶斯滤波问题的算法,它通过先前时间阶估计得到的后验密度迭代产生提议分布。序贯重要性采样算法的缺点



在于粒子退化问题[71]，也就是说，一旦提议分布与实际滤波分布的形状匹配不佳，大部分粒子的权重将变得很小，这些粒子在传播过程中其影响将逐渐趋于零，这将导致大量的计算代价浪费在低权重的粒子上。另一方面，因子采样算法的缺陷在于样本贫化现象，即多个样本落在样本空间中的同一点，导致样本集缺乏多样性。

通过在状态空间中维持多个假设，粒子滤波算法能够比较有效地处理多模分布问题。作为一个一般性的通用跟踪框架，尽管它在视觉跟踪领域取得了很大成功，但是粒子滤波算法本身在处理高维状态空间的跟踪问题时仍然具有局限性。为了克服这一局限性，一些研究人员尝试通过多种途径来进一步改进粒子采样效率。近年来，将密度估计方法引入到粒子滤波框架成为了提高粒子采样效率的有效途径之一[76~78]。Cham 和 Rehg[79]使用一个分片的高斯模型来表示观测似然度量和先验预测密度，从而获得后验密度的解析计算方式，实现多假设跟踪。Vermaak 等人[80]则直接将后验密度建模为高斯混合模型，作者指出后验密度的传播可以通过对混合模型的单个成分独立地进行滤波，每个成分都获得一定数目的粒子，以此实现多假设跟踪。Han 等人[76]将核密度估计技术引入到序贯贝叶斯滤波框架实现快速的后验密度解析和估计。此外，密度估计的支持向量方法也在序贯贝叶斯滤波中得到了应用[81]。Baner jee 和 Burlina[81]采用稀疏核密度估计的方法计算给定粒子集的先验密度，其中的密度估计器包括支持向量描述（SVDD）和相关向量机（RVM）[82]。在粒子滤波框架中采用支持向量的方法估计密度，其目的是获得后验密度的稀疏表示和解析解，从而降低粒子滤波的计算代价。

基于密度估计的技术应用在贝叶斯滤波有助于简化采样问题，同时在采样过程中也可以维持多个假设，但是粒子滤波算法仍然受限于高昂的计算代价，尤其是在复杂滤波分布条件下需要使用大量的粒子来近似后验分布。为了降低计算负担提高粒子采样效率，学者们提出了多种粒子滤波改进算法。Deutcher 等人[83]在传统的粒子滤波跟踪算法[4]中交错地使用一个模拟退火过程。模拟退火过程便于从平滑化的后验分布采样，这种结合使得抽取的样本靠近实际的后验模式，从而避免采样陷入局部最优模式。该方法的一个局限性是由于在采样中调用模拟退火算法，缓慢的退火过程使得滤波算法的计算复杂度仍然很高。Choo 和 Fleet[84]提出在序贯贝叶斯框架下使用 MCMC 采样方法替代广泛采用的重要性采样算法，直接从后验分布采样。这种采样方式主要是通过一个合适的提议分布从后验分布中采样，以此构建一条收敛于后验分布的马氏链。相比与传统的序贯重要性采样算法，使用 MCMC 采样的好处在于可以更高效地从后验分布中采样，因而可以使

用较少的粒子来近似目标分布。与此工作类似，Khan 等人[5]提出一种基于 MCMC 采样的粒子滤波算法用于多目标跟踪。作者直接在高维的状态空间中建模多目标状态，通过引入一个基于马尔科夫场的运动先验来描述多目标跟踪场景下目标间的相互影响，实现在高维状态空间中的 MCMC 采样。他们的实验表明，相比于独立的多个粒子滤波实现，使用 MCMC 采样替代序贯重要性采样可以显著地减少所需的粒子数目从而降低计算复杂度。紧接着，Zhao 等人[85]提出采用数据驱动的 MCMC 采样思路替代传统的 MCMC 方法，在采样过程中更有效地估计提议分布，使产生的样本更靠近样本空间中的局部模式，从而提高采样效率、降低计算复杂度。

尽管 SMC 采样方法和多种 MCMC 采样方法在贝叶斯滤波跟踪问题中表现出良好的可扩展性，但是对于复杂多模的滤波分布，现有的随机跟踪方法在采样效率方面还有待进一步提高。探索更为高效的贝叶斯滤波跟踪方法仍然是个极具挑战性的研究课题。

3.3 方法介绍

3.3.1 贝叶斯滤波

在贝叶斯滤波跟踪框架下，视觉跟踪可看作状态空间模型中系统的状态估计问题。在此框架下，目标的状态是一个随机变量，目标跟踪的任务是基于系统的动态模型，从噪声观测数据中估计后验概率（滤波）密度。系统定义为

$$\begin{cases} x_t = f(x_{t-1}, u_t) \\ z_t = h(x_t, v_t) \end{cases} \tag{3.1}$$

式中，$x_t \in \mathbb{R}^n$ 表示目标在 t 时刻的状态，是要求解的状态变量；函数 f 为系统的动态模型，定义了系统状态随时间的演化规律，它通常定义为一阶马尔科夫模型；u_t 表示系统噪声影响；$z_t \in \mathbb{R}^m$ 表示系统的图像观测；函数 h 为系统的观测模型，描述了目标状态与图像观测的关联；v_t 为观测噪声。

在贝叶斯滤波框架下，系统的动态模型和观测模型可以分别用概率分布 $p(x_t | x_{t-1})$ 和 $p(z_t | x_t)$ 来表示。按照贝叶斯律，可以在各个时刻 t 序贯地估计目标状态 x_t 的后验密度 $p(x_t | z_{1:t})$。图 3-1 所示为状态与观测在图模型（Graphical Model）下的表示。



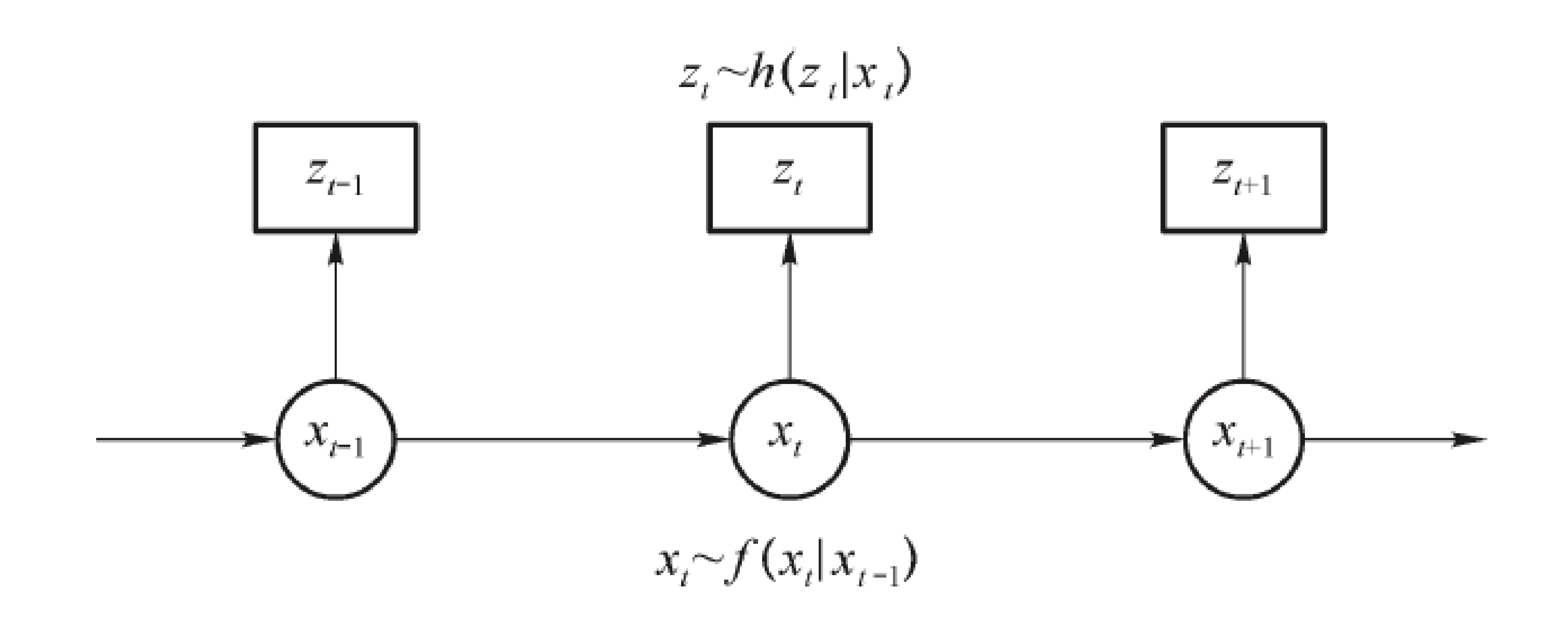


图 3-1　状态与观测在图模型下的表示

记 $z_{1:t}=\{z_1,z_2,\cdots,z_t\}$ 为从初始到当前时刻 t 的观测集，贝叶斯滤波的实现依赖于两个迭代进行的阶段，即预测和更新。

预测：
$$p(x_t \mid z_{1:t-1}) = \int p(x_t \mid x_{t-1} p)(x_{t-1} \mid z_{1:t-1})\mathrm{d}x_{t-1} \tag{3.2}$$

更新：
$$p(x_t \mid z_{1:t} \propto p(z_t \mid x_t)p(x_t \mid z_{1:t-1}) \tag{3.3}$$

预测阶段使用传播密度 $p(x_t|x_{t-1})$ 计算先验分布，而更新阶段使用观测度量 $p(z_t|x_t)$ 更新后验。当噪声为高斯分布且系统的演化和度量均为线性时，卡尔曼滤波可以获得精确式(3.1)的解析解。然而在实际应用中，系统多为非高斯条件下的非线性系统。在这种情况下，以高斯噪声作为基本假设的卡尔曼滤波，扩展的卡尔曼滤波(EKF)，以及 Unscented 卡尔曼滤波(UKF)均无法很好地解决贝叶斯滤波问题。

3.3.2　粒子滤波

粒子滤波[4,71]，又称序贯蒙特卡洛方法，是实现贝叶斯滤波的一种 MC 采样技术。它在每个时间阶 t 使用一个加权的粒子集 $\{x_t^{(i)},w_t^{(i)}\}_{i=1}^N$ 来近似后验密度。各个时间阶的粒子集通过系统的动态模型传播到下一个时间阶。通过这些粒子和权重，可以获得目标状态的估计值。粒子滤波的实现有多种版本，从采样的角度看粒子滤波算法可以归为三类，即重要性采样、因子采样以及拒绝采样。这里我们仅描述最为一般的算法实现，它包含了不同版本的粒子滤波算法的主要特征，容易推导出不同版本的粒子滤波算法实现。

算法输入：粒子集 $\{x_{t-1}^{(i)},w_{t-1}^{(i)}\}_{i=1}^N$，观测 z_t；

算法输出：t 时刻粒子集 $\{x_t^{(i)},w_t^{(i)}\}_{i=1}^N$，状态估计 $\hat{x}_t$；

(1) 从提议分布中抽取样本：

$$x_t^{(i)} \sim q(\cdot \mid x_{t-1}^{(i)},z_t),i=1,2,\cdots,N$$

（2）粒子加权：

$$w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(z_t \mid x_t^{(i)})p(x_t^{(i)} \mid x_{t-1}^{(i)})}{q(x_t^{(i)} \mid x_{t-1}^{(i)}, z_t)}, i = 1,2,\cdots,N$$

（3）归一化权重：

$$w_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(j)}}, i = 1,2,\cdots,N$$

（4）计算粒子集的有效样本量 N_{ESS}：

$$N_{\text{ESS}} = \sum_{i=1}^{N} (w_t^{(i)})^2$$

（5）If $N_{\text{eff}} < N_0$ then //重采样

– 估计累积分布 f：

$$f_i = \sum_{j=1}^{i-1} f_j + w_t^{(i)}, i = 1,2,\cdots,N$$

– For $i=1$ to N

– $u_i \sim Unif(0, f_n)$

– 求 k：$(u_i - f_k)(u_i - f_{k+1}) \leqslant 0$

– $x_t^{(i)} = x_t^{(k)}$；$w_t^{(i)} = \dfrac{1}{N}$

– End For

End If

（6）状态估计：

$$\hat{x}_t = \sum_{i=1}^{N} w_t^{(i)} x_t^{(i)}$$

这里描述的粒子滤波算法容易推导出常用的 SMC 算法[71]。如果不考虑重采样步骤(5)，算法等价于序贯重要性采样(SIS)算法；如果重采样步骤(5)不依赖于有效样本大小 N_{ESS} 且总是包含在算法中，那么该算法就等价于著名的 Consendation 算法[4]，或者采样重要性重采样(SIR)算法。类似地，如果密度传播不考虑当前的观测度量，即 $p(x_t | x_{t-1}, z_t) = p(x_t | x_{t-1})$，那么算法等价于 Bootstrap 粒子滤波[71]。关于有效样本大小 N_{ESS} 的讨论，参见本书第 2 章的重要性采样部分。

如图 3-2 所示，我们以经典的 SIR 算法为例，分析粒子滤波算法的一个迭代过程[143]。给定在 $t-1$ 时刻的等权重粒子集 $\left\{\tilde{x}_{t-1}^{(i)}, \frac{1}{N}\right\}_{i=1}^{N}$，它用于近似分布 $p(x_{t-1} | z_{1:t-2})$。在 $t-1$ 时刻，使用观测度量对粒子进行加权 $\{\tilde{x}_{t-1}^{(i)}, \tilde{w}_{t-1}^{(i)}\}_{i=1}^{N}$，它用于近似



$t-1$ 时刻的后验分布 $p(x_{t-1}|z_{1:t-1})$。然后，执行重采样步骤获得等权重粒子集 $\left\{x_{t-1}^{(i)},\frac{1}{N}\right\}_{i=1}^{N}$，此时粒子集仍然代表后验分布 $p(x_{t-1}|z_{1:t-1})$。最后，粒子集通过提议分布（重要性函数）传播到 t 时刻生成 $\left\{\tilde{x}_t^{(i)},\frac{1}{N}\right\}_{i=1}^{N}$，用以近似分布 $p(x_t|z_{1:t-1})$。至此，完成了算法从 $t-1$ 时刻到 t 时刻的一次迭代过程。

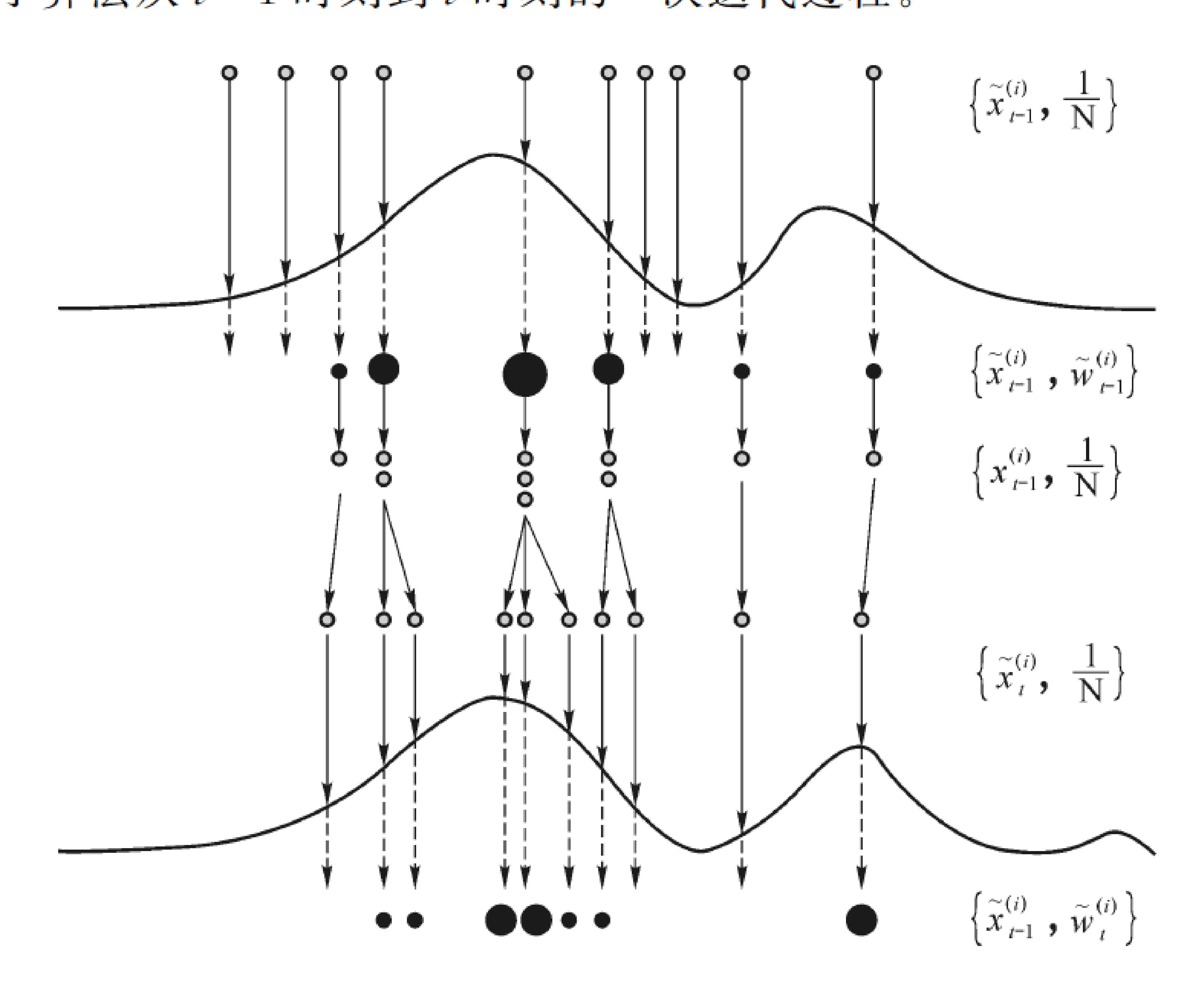


图 3-2　采样重要性重采样(SIR)算法的一次迭代过程。

在不同版本的粒子滤波实现中，Bootstrap 粒子滤波[71]是较为吸引人的一种算法。它直接使用先验分布（动态模型）作为提议分布，从而将粒子加权步骤(2)简化为

$$w_t^{(i)} \propto w_{t-1}^{(i)} p(z_t \mid x_t^{(i)}), i = 1,2,\cdots,N \tag{3.4}$$

这意味着粒子权重的更新只依赖于当前的观测密度 $p(z_t|x_t)$，使得采样和计算过程更为简化和直观。尽管采用先验密度作为提议分布能够给滤波计算带来简化，但是也容易导致滤波采样性能的下降。其主要原因在于从观察序列中获取的信息往往带有噪声，缺乏稳健性，尤其是当先验分布与基于当前观测的后验分布之间匹配不佳（支撑区域的交集较小）时。一种可行的方案是在提议分布中既考虑粒子的历史状态也兼顾当前的观测信息[144]，例如，使用如下的提议分布：

$$q(x_t \mid x_{t-1}, z_t) = \frac{p(x_t \mid x_{t-1})p(z_t \mid x_t)}{\int p(x \mid x_{t-1})p(z_t \mid x)\mathrm{d}x} \tag{3.5}$$

这就是所谓的“最优”提议分布[144]。其主要特征是，给定粒子的历史轨迹，粒子权重的条件方差为零。基于这一提议分布，粒子加权按下式进行：

$$w_t^{(i)} \propto w_{t-1}^{(i)} p(z_t \mid x_{t-1}) = w_{t-1}^{(i)} \int p(x \mid x_{t-1}^{(i)}) p(z_t \mid x) \mathrm{d}x \tag{3.6}$$

“最优”提议分布产生的粒子倾向于聚集在条件后验密度很高的区域，不同于采用动态模型作为提议分布时在状态空间中盲目地产生粒子。尽管这种提议分布在理论上具有优良的性质，但式(3.6)中的积分项在多数情况下无解析形式，因而需要采用一些近似技术来计算，包括使用非线性滤波的局部线性化、高斯近似等方法[71]。

3.3.3　基于 MCMC 采样的粒子滤波

尽管粒子滤波由于其良好的可扩展性在视觉跟踪领域得到了广泛的应用，但是算法实现仍然受限于高昂的计算代价，尤其是在复杂滤波分布条件下需要大量的粒子来近似后验分布。Choo 和 Fleet[84]提出在序贯贝叶斯框架下使用 MCMC 采样方法替代广泛采用的重要性采样算法，直接从后验分布采样。这种采样方式主要是通过一个合适的提议分布从后验分布中采样，以此构建一条收敛于后验分布的马氏链。相比传统的序贯重要性采样算法，使用 MCMC 采样的好处在于可以更高效地从后验分布中采样，因而可以使用较少的粒子来近似目标分布。与此工作类似，Khan 等人[5]提出一种基于 MCMC 采样的粒子滤波算法用于多目标跟踪。这些研究工作表明，使用 MCMC 采样替代序贯重要性采样可以显著地减少所需的粒子数目，从而降低计算复杂度。

基于 MCMC 采样的粒子滤波算法(MCMC-PF)使用 Metropolis-Hastings 算法构建一条收敛于贝叶斯滤波分布的马氏链。假定 $t-1$ 时刻获得的粒子集为 $\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$，则滤波分布 $p(x_t \mid z_{1:t})$ 可以近似地表示为

$$p(x_t \mid z_{1:t}) \propto p(z_t \mid x_t) \sum_{i=1}^{N} w_{t-1}^{(i)} p(x_t \mid x_{t-1}^{(i)}) \tag{3.7}$$

假定提议分布为 q，在第 i 次迭代链的状态为 $x_t^{(i)}$，MCMC-PF 使用提议分布 q 产生一个候选样本 x'，并以概率 $\alpha(x_t^{(i)}, x')$：

$$\alpha(x_t^{(i)}, x') = \min\left\{1, \frac{p(x' \mid z_{1:t}) q(x_t^{(i)}; x')}{p(x_t^{(i)} \mid z_{1:t}) q(x'; x_t^{(i)})}\right\} \tag{3.8}$$

接受 x' 为马氏链的新样本，否则保留 $x^{(i)}{}_t$。

值得注意的是，MCMC-PF 使用 Metropolis-Hastings 算法直接从后验分布中



采样，因而构建的马氏链以滤波分布为极限平稳分布，这意味着获得的样本集是等权重或无权重的(Unweighted)，这与重要性采样有着本质的不同。于是式(3.7)可以改写为

$$p(x_t \mid z_{1:t}) \propto p(z_t \mid x_t)\sum_{i=1}^{N} p(x_t \mid x_{t-1}^{(i)}) \tag{3.9}$$

采样过程结束后，对状态变量的估计可通过马氏链上的样本点采用 MSE 准则或 MAP 准则进行。状态的 MAP 估计为

$$x_t^{\mathrm{MAP}} = \mathrm{argman}_{x_t} p(x_t \mid z_{1:t}), x_t \in \{x_t^{(i)}\}_{i=1}^{N} \tag{3.10}$$

如果状态估计采用 MSE 准则，一般需要对马氏链上一定数量的初期样本做丢弃处理(Burn-In)，以降低估计误差。关于马氏链的收敛诊断与分析详见本书的第 2 章。

MCMC-PF 算法描述如下：

算法输入：粒子集$\{x_{t-1}^{(i)}\}_{i=1}^{N}$，观测 z_t；

算法输出：粒子集$\{x_t^{(i)}\}_{i=1}^{N}$，状态估计 x_t^{MAP}；

(1) 从粒子集$\{x_{t-1}^{(i)}\}_{t-1}^{N}$中随机抽取一样本 $x_{t-1}^{(j)}$；

(2) 产生初始样本 $x_t^{(0)} \sim p(x_t \mid x_{t-1}^{(j)})$；

(3) For $i=1$ to N

- 抽取候选样本 $x' \sim q(\cdot\ ; x_t^{(i-1)})$；
- 按照式(3.8)计算接受比 α；
- 以概率 α 接受 x'，并令 $x_t^{(i)} = x'$；否则令 $x_t^{(i)} = x_t^{(i-1)}$；

End For

(4) 按照式(3.10)估计 x_t^{MAP}.

MCMC-PF 算法尽管能够在一定程度上克服粒子滤波算法在采样过程中存在的粒子贫化现象，但是传统的 MCMC 算法，例如 Metropolis-Hastings 算法，在对复杂多模的滤波分布进行采样时往往表现出混合性差，收敛缓慢等特点，使得采样可能长时间停留在某个状态无法从局部模式中跳出。此外，MCMC-PF 算法中使用的提议分布一般采用局部高斯扰动，因而构建的马氏链事实上是一条随机游走链。提议分布延展度的设置往往很难把握，过大或过小的方差都可能造成采样性能的下降。这些局限性使得 MCMC-PF 算法在处理具有复杂多模分布的视觉跟踪问题时，往往容易导致目标跟踪的失败。

3.3.4　序贯随机逼近蒙特卡洛跟踪方法

通过将随机逼近蒙特卡洛(SAMC)采样方法[93]引入到贝叶斯滤波跟踪框架

中，本章提出一种序贯随机逼近蒙特卡洛(Sequential SAMC)跟踪方法。它能显著地提高 MCMC 算法对滤波分布的采样效率，同时能够有效地避免 MCMC 采样陷入局部模式。本节首先分析 MCMC 算法应用于滤波分布采样时易于陷入局部模式的原因，然后给出我们的解决方案，最后介绍 Sequential SAMC 跟踪算法的实现细节。

正如 MCMC-PF 算法所介绍的，MCMC 方法可替代常用的 SIR 算法对滤波分布进行采样，以获得更为理想的近似。给定目标分布(即滤波分布)$p(x_t|z_{1:t})$和当前链的状态 x_t，一个 Metropolis-Hastings 采样步骤包括：使用提议分布 $q(\cdot;x_t)$抽取一个候选样本 x'，马氏链以概率 $\alpha(x';x_t)$从状态 x_t 跳跃到 x'，其中

$$\alpha(x';x_t) = \min\left\{1, \frac{p(x' \mid z_{1:t})q(x_t;x')}{p(x_t \mid z_{1:t})q(x';x_t)}\right\} \tag{3.11}$$

从式(3.11)可以看到，如果链的当前状态处于滤波分布的峰值附近，或提议的样本处于滤波分布的尾部，那么采样过程就可能陷入局部模式，很难从中逃离并跳跃到别的模式。对于复杂多模的滤波分布(例如，场景中存在多个相似的运动目标，或目标运动具有极大不确定性等因素)来说，其能量分布往往表现出崎岖不平的特征，MCMC 采样能否有效地跨越高能量屏障到达样本空间的其他模式，决定了滤波跟踪的输出性能。从能量空间的角度看，如果采样算法能够在滤波分布的各个能量阶子区域进行随机游走，那么就可以从本质上有效地避免采样陷入局部模式。

记$\mathcal{X}$为滤波分布 $p(x_t|z_{1:t})$的样本空间。首先按照能量阶(Level)把样本空间$\mathcal{X}$划分为 m 个不相交的子区域(Subregion)，即$\mathcal{X}=\bigcup\mathcal{X}_i$，$i=1,2,\cdots,m$。特别地，按照实值能量函数 $E(x_t)=-\log\{p(x_t|z_{1:t})\}$把样本空间$\mathcal{X}$划分为 m 个等能量带(Energy Band)：

$$\begin{aligned}
\mathcal{X}_1 &= \{x: E(x)\leqslant u_1\}\\
\mathcal{X}_2 &= \{x: u_1 < E(x)\leqslant u_2\}\\
&\vdots\\
\mathcal{X}_{m-1} &= \{x: u_{m-2} < E(x)\leqslant u_{m-1}\}\\
\mathcal{X}_m &= \{x: E(x) > u_{m-1}\}
\end{aligned}$$

式中，$u_1,u_2,\cdots u_{m-1}$为 $m-1$ 个实数。

接着，我们构造一个试验分布：

$$p_g(x_t) \propto \sum_{i=1}^{m} \frac{p(x_t \mid z_{1:t})}{g_i} I(x_t \in \mathcal{X}_i) \tag{3.12}$$

其中，I 为示性函数；$g_i = \int_{\mathcal{X}_i} p(x_t \mid z_{1:t})\mathrm{d}x_t$；在物理学中，$g=\{g_i,i=1,2,\cdots,m\}$可



以看作连续分布系统下的态密度(DoS)项。直观上看，如果态密度 g 在采样过程中能够被准确地估计，那么从试验分布 $p_g(x_t)$ 抽样意味着马氏链将会在能量空间进行随机游走，从而采样的局部陷入问题将会得到有效地解决。

构造试验分布式(3.12)替代原始的滤波分布作为目标分布并从中采样，其用意就是在于通过 MCMC 采样自适应地估计目标分布，通过对态密度项的估计来间接地抽取服从滤波分布的(加权)样本。这种采样思路在一定程度上是受到 Multicanonical 系统采样算法及其变体[91~93]的启发。正如本书第 2 章所介绍的，Multicanonical 采样算法及其多种变体在物理学领域应用广泛，通常用于对一些物理量的估计和计算，它能够有效地避免在采样过程中陷入局部能量极小点。这些采样算法大多面向全局优化，或统计量估计。然而，对于贝叶斯滤波跟踪问题，我们既要考虑如何避免采样陷入局部模式，还需要考虑粒子传播、粒子加权等问题。接下来我们详细阐述如何将随机逼近蒙特卡洛采样方法引入到贝叶斯滤波框架，提出一种序贯随机逼近蒙特卡洛跟踪方法，实现更为高效、稳健的视觉跟踪。

1. 随机逼近蒙特卡洛采样

随机逼近蒙特卡洛(SAMC)算法是一种基于目标分布自适应的 MCMC 算法，同时也属于 Robbins-Monro[115]类型的随机逼近算法。它通过在 MCMC 算法中嵌入一个随机逼近过程(SAP)，实现对目标分布的自适应更新与估计，从而在采样过程中持续地改进统计量(例如态密度)的估计精度。同时，它具有 Multicanonical 采样算法的优点，能够有效地避免采样陷入局部模式。

假定目标分布为贝叶斯滤波分布 $p(x_t|z_{1:t})$。一次 SAMC 采样过程描述如下：

(1) Metropolis-Hastings 采样

– 给定当前链上的样本 $x_t^{(k)}$，使用提议分布 $q(\cdot;x_t^{(k)})$ 抽取一个候选样本 x'；

– 计算 Metropolis-Hastings 接受比 α：

$$\alpha(x';x_t^{(k)}) = \min\left\{1,\frac{p_g(x')q(x_t^{(k)};x')}{p_g(x_t^{(k)})q(x';x_t^{(k)})}\right\} \tag{3.13}$$

– 以概率 α 接受 x' 为新样本，令 $x_t^{(k+1)}=x'$；

(2) 工作权重(working weight)更新

$$\theta_i^{(k+1)} = \theta_i^{(k)} + \gamma_k[I(x_t^{(k+1)} \in \mathcal{X}_i) - \phi_i],i = 1,2,\cdots,m \tag{3.14}$$

式中，$\theta_i=\log g_i$；$\phi=\{\phi_i,i=1,2\cdots,m\}$，表示在各个能量子区域上的期望采样分布；$\gamma_k$ 为增益因子(Gain Factor)，它控制工作权重的学习速度。在 Robbins-Monro 型的随机逼近算法中，要求增益因子满足如下条件：

$$\lim_{k\to\infty}\gamma_k = 0; \quad \sum_{k=1}^{\infty}\gamma_k = \infty; \quad \sum_{k=1}^{\infty}\gamma_k^{\eta} < \infty \tag{3.15}$$

式中，$\eta\in(1,2)$。在 SAMC 中，Liang 等人[93]使用如下的学习率：

$$\gamma_k = \frac{k_0}{\max(k_0, k)} \tag{3.16}$$

式中，$k_0>0$，它控制着学习率。

在贝叶斯滤波中，在任意时刻 t，我们可以使用 SAMC 算法从后验分布 $p(x_t|z_{1:t})$中采样，在获得足够数量的样本集后再对 x_t 的统计量进行估计。对应于贝叶斯滤波跟踪问题，就是通过获取的样本集按照 MSE 准则或 MAP 准则对目标状态进行估计。

值得注意的是，在贝叶斯滤波框架下，对后验分布的估计依赖于先前时刻的样本历史轨迹，因而我们需要考虑粒子的随机传播、粒子加权等问题。接下来我们详细介绍如何使用 SAMC 采样方法构建一个动态加权的重要性采样（Dynamical Weighted Importance Sampling，DWIS）算法，即 Sequential SAMC 算法，应用于贝叶斯滤波跟踪。

2. 序贯随机逼近蒙特卡洛

与基本的粒子滤波算法和 MCMC-PF 算法不同，本节提出的序贯随机逼近蒙特卡洛（Sequential SAMC，S-SAMC）算法包含三个主要的步骤：条件密度传播、SAMC 采样和样本加权（Reweighting）。采样算法流程如图 3-3 所示。

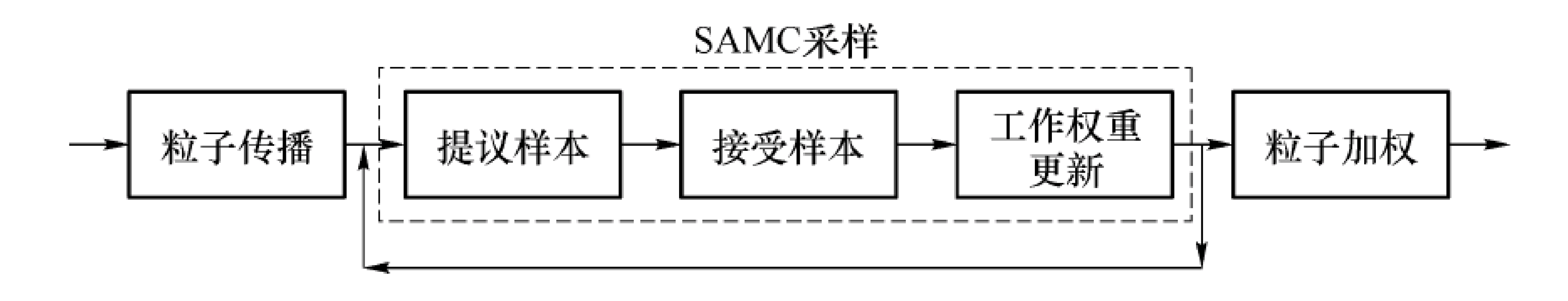


图 3-3　S-SAMC 采样算法流程

（1）条件密度传播。该阶段也称为粒子传播，其目的是传播前一时间阶获得的粒子集到当前时间阶，用于估计当前时间阶的后验密度。给定前一时间阶的粒子集 $\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$，首先计算其有效样本大小 N_{ESS}，若小于一特定的阈值则执行一次重采样操作。基于此粒子集，结合先验模型 $p(x_t|x_{t-1})$与当前时间阶的观测度量 $\mathrm{p(z_t|x_t)}$，当前时间阶的后验密度（滤波密度）$\mathrm{p(x_t|z_{1:t})}$可近似为

$$p(x_t \mid z_{1:t}) \propto p(z_t \mid x_t)\sum_{i=1}^{N} w_{t-1}^{(i)} p(x_t \mid x_{t-1}^{(i)}). \tag{3.17}$$

此外，鉴于我们的目的是构建一条平稳收敛于滤波分布的马氏链，因而需要产



生一个随机初始样本 $x_t^{(0)}$。初始样本的产生有多种方式。一种典型方式是从 $\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$ 中抽取一样本 $x_{t-1}^{(j)}$，然后使用先验密度传播到 $x_t^{(j)}$ 作为初始样本。另一种可行方案是使用先前时间阶的目标状态 MAP 估计，通过二阶自回归模型(AR2)产生当前时间阶的初始样本。无论采用哪种方式，构建的马氏链都将收敛于平稳分布 $p(x_t|z_{1:t})$。事实上，SAMC 采样对初始样本的状态依赖较小，考虑使用 AR2 模型或先验密度产生初始样本是基于目标运动连续性的假设。当目标运动动力学表现出极大不确定性时，Sequential SAMC 也能在采样过程中覆盖这种不确定性。这里采用 AR2 模型确定 $x_t^{(0)}$，即

$$x_t^{(0)} \sim \mathcal{N}(\mu, \Sigma) \tag{3.18}$$

其中，$\mathcal{N}(\mu, \Sigma)$ 表示均值为 $\mu = A_{1x_{t-1}^*} + A_{2x_{t-2}^*}$ 方差为 Σ 的高斯分布，x_{t-1}^* 和 x_{t-2}^* 为先前时间阶目标状态的 MAP 估计，A_1 和 A_2 定义一个常加速度的 AR2 模型。

(2) SAMC 采样。SAMC 采样阶段的任务是使用前面介绍的 SAMC 算法构建一条收敛于目标分布的马氏链。采样过程包含三个步骤：样本提议、样本接受和工作权重更新。

① 样本提议阶段。此阶段的任务是设计和选择一个提议分布 q，用于产生候选样本。我们考虑一个局部高斯扰动的提议分布：

$$q(\cdot\,; x_t^{(k)}) \propto \mathcal{N}(\cdot\,; \mu, \Sigma) \tag{3.19}$$

其中，均值 μ 为当前链的样本状态 $x_t^{(k)}$，易于验证该提议分布满足提议的全局性(Global)要求[93]，即对于任意的 $x, x' \in \mathcal{X}$，都有 $q(x'; x) > 0$。全局性提议分布是确保基于目标分布自适应的 SAMC 算法的必要条件之一。构建其他符合全局性要求的提议分布也是可行的，但这种基于局部扰动的提议是最为简单且易于实现的方案之一。在第 4 章中我们将介绍更为复杂的提议设计方法，用以处理大范围运动突变的视觉跟踪问题。

② 样本接受阶段。给定马氏链在第 k 次迭代时的状态 $x_t^{(k)}$，以及提议 q 产生的候选后续样本 x'，样本接受取决于 Metropolis-Hastsings 接受比 $\alpha(x'; x_t^{(k)})$。值得注意的是，此时的目标分布为试验分布 $p_g(x_t)$，而不是原始的滤波分布。

$$\alpha(x'; x_t^{(k)}) = \min\left\{1, \frac{p(x' \mid z_{1:t})}{p(x_t^{(k)} \mid z_{1:t})} \frac{g_{J(x_t^{(k)})}^{(k)}}{g_{J(x')}^{(k)}} \frac{q(x'; x_t^{(k)})}{q(x_t^{(k)}; x')}\right\} \tag{3.20}$$

式中，$J(x)$ 代表样本 x 所属的能量子区域的索引，即若 $x \in \mathcal{X}_j$，则 $J(x) = j$。

③ 工作权重更新阶段。一旦采样算法产生一个新的样本，例如 x'，就按照式(3.14)更新工作权重 θ。这意味着在采样过程中态密度项的估计依赖于马氏链的历史样本信息，并通过一个随机逼近过程来迭代学习的。相比于 Wang-Landau 算



法，使用式(3.14)的随机逼近算法对态密度的估计更为有效。随着采样的进行，SAP 能够持续地改进态密度的估计精度。通过式(3.14)和式(3.19)我们可以看到，一旦一个候选样本(例如 x')被拒绝，马氏链停留在原地且当前样本状态值被保留，同时，当前样本所属能量区域的态密度值被调高，从而在下次采样时从当前样本状态跳出的概率增大。与此同时，除当前样本所属的能量区域外，其他能量子区域的态密度值被调低，使得下次采样时马氏链跳跃到这些区域的机会变得更大。这种工作权重(即态密度)的自校正机制使得对滤波分布的采样不易于陷入局部模式，从而提高在复杂多模情况下的采样混合性，使得通过少量粒子达到或靠近全局最优模式成为可能。

(3) 粒子加权。回顾粒子滤波算法，包括 SIR[74]、Bootstrap 粒子滤波[71]和 Condensation[4]等，每个粒子在产生之后通过特定的加权度量立即被赋予相应的权重，在后续的采样过程中粒子权重始终保持不变。而在 MCMC-PF 中，由于 MCMC 算法直接从后验分布中采样，产生的粒子集以滤波分布为平稳分布，因而无须加权操作。

本章提出的 SAMC 算法有别于以上这两类算法。尽管它也是在贝叶斯滤波框架下使用 MCMC 采样算法(SAMC)，但由于是从一个试验分布中采样，因而产生的样本集以试验分布为平稳目标分布。按照重要性采样理论，样本集通过重要性加权后将近似地服从原始的滤波分布。此外，S-SAMC 算法属于动态重要性采样算法的范畴[97,98]，这意味着在构建一条自适应马氏链的过程中，样本的权重是动态估计的。注意到在 SAMC 采样过程中式(3.14)(SAP)的任务正是在线学习这个动态权重，即态密度项。使用一个 SAP 在线估计样本权重的好处在于，利用在采样过程中收集的历史样本信息指导采样向高后验密度区域偏置，使对样本权重的估计可以持续地被改进。于是，当采样结束并获得样本集 $\{x_t^{(i)}\}_{i=1}^N \sim p_g(x_t)$ 之后，可通过下式对粒子进行加权：

$$w_t^{(i)} = g_{J(x_t^{(i)})}, i = 1,2,\cdots,N \tag{3.21}$$

使得 $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^N \sim p(x_t | z_{1:t})$，其中，$J(x)$ 代表样本 x 所属的能量子区域的索引，即若 $x \in \mathcal{X}_j$，则 $J(x) = j$。在使用 S-SAMC 算法获得粒子集 $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^N$ 后，可通过下式获得目标状态的期望估计：

$$\hat{x}_t = \frac{\sum_{i=1}^N w_t^{(i)} x_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}} \tag{3.22}$$



事实上，对于动态重要性采样算法，有如下的定理[145]：

定理 3.1　假定使用 SAMC 算法从 $f(x)$ 产生样本集 $\{x^{(i)},w^{(i)}\}_{i=1}^{n}$，并满足条件式(3.15)和(A.1)，那么对任一可积函数 $h(x)$，它在分布 f 下的期望 $E_f h(x)$ 可通过下式进行估计：

$$\widehat{E_f h(x)} = \frac{\sum_{i=1}^{n} w^{(i)} h(x^{(i)})}{\sum_{i=1}^{n} w^{(i)}} \tag{3.23}$$

当 $n \to \infty$，$\widehat{E_f h(x)} \to E_f h(x)$。其中条件(A.1)表述如下：

(A.1) 记 K_θ 为一个 Metropolis-Hastings 转移核，其中 $\theta \in \Theta$。对于任意的 $\theta \in \Theta$，K_θ 是非周期不可约的。此外，存在一个常数 $\alpha \leqslant 2$ 和函数 $V:\mathcal{X} \to [1,\infty)$，对于任意一紧集 $\mathcal{K} \in \Theta$，都有：

(a) 存在一个整数 l，常数 $0<\lambda<1$，$b,\xi,\sigma>0$，$\mathcal{L} \in \mathcal{X}$ 和一概率度量 v，使得

$$\begin{cases} \sup_{\theta\in\mathcal{K}} K_\theta^l V^\alpha(x) \leqslant \lambda V^\alpha(x) + bI(x \in \mathcal{L}),\ \forall x \in \mathcal{X}; \\ \sup_{\theta\in\mathcal{K}} K_\theta V^\alpha(x) \leqslant \xi V^\alpha(x),\ \forall x \in \mathcal{X}; \\ \inf_{\theta\in\mathcal{K}} K_\theta^l(x,A),\ \forall x \in \mathcal{L},\ \forall A \in \mathcal{B}x; \end{cases}$$

其中，$\mathcal{B}\mathcal{X}$ 是定义于样本空间 $\mathcal{X}$ 上的 Borel 集，$K_\theta V(x) = \int_{\mathcal{X}} K_\theta(x,y)V(y)\mathrm{d}y$。

(b) 存在一常数 c 使得对任意 $(\theta,\theta') \in \mathcal{K}\times\mathcal{K}$ 都有

$$\begin{cases} \|K_\theta g - K_{\theta'} g\|_V \leqslant c\|g\|_V \mid \theta-\theta' \mid,\ \forall g \in \mathcal{L}_V; \\ \|K_\theta g - K_{\theta'} g\|_{V^\alpha} \leqslant c\|g\|_{V^\alpha} \mid \theta-\theta' \mid,\ \forall g \in \mathcal{L}_{V^\alpha}; \end{cases}$$

式中，$\|g\|_V = \sup_{x\in\mathcal{X}} |g(x)/V(x)|$；$\mathcal{L}_V = \{g:\mathcal{X} \to \mathbb{R}^m, \|g\|_V < \infty\}$。

由定理 3.1 易于验证当 $N \to \infty$，由式(3.21)估计的 $\hat{x}_t$ 收敛于目标状态的期望。综合以上分析，S-SAMC 采样算法描述如下：

算法输入：粒子集 $\{x_{t-1}^{(i)},w_{t-1}^{(i)}\}_{i=1}^{N}$；

算法输出：粒子集 $\{x_t^{(i)},w_t^{(i)}\}_{i=1}^{N}$，状态估计 $\hat{x}_t$。

(1) 初始化

- 样本空间划分：$\mathcal{X} = \{\mathcal{X}_i\}_{i=1}^{m}$
- 初始化工作权重：$g_i = 1, i = 1,2,\cdots,m$

(2) 重采样 $\{x_{t-1}^{(i)},w_{t-1}^{(i)}\}_{i=1}^{N}$，产生等权重粒子集 $\left\{x_{t-1}^{(i)},\frac{1}{N}\right\}_{i=1}^{N}$；

(3) 按照式(3.19)产生初始样本 $x_t^{(0)}$；

(4) For $i=1$ to N

- 按式(3.19)抽取候选样本 $x' \sim q(\cdot, x_t^{(i-1)})$
- 按照式(3.20)计算接受比 α
- 以概率 α 接受 x'，并令 $x_t^{(i)}=x'$；否则令 $x_t^{(i)}=x_t^{(i-1)}$
- 按式(3.14)更新工作权重

End For

(5) 粒子加权

- 按式(3.21)对样本集加权
- 归一化粒子集

(6) 按式(3.22)估计目标状态的期望 $\hat{x}_t$

3. S-SAMC 跟踪方法

由前面的算法分析可以看到，本章提出的 S-SAMC 算法与粒子滤波、MCMC-PF 等算法类似，都是对贝叶斯滤波问题的一种算法实现。因而当 S-SAMC 采样算法应用于视觉跟踪时，它是一个较为通用的一个视觉跟踪框架。它能够依照不同的跟踪应用容纳不同的观察模型、先验模型和目标表示。本节讨论基于 S-SAMC 采样的视觉跟踪方法，涉及目标表示、运动先验、观测模型和参数设置等方面的内容。

考虑二维的视觉跟踪，目标采用简单的矩形框描述，状态变量定义为 $x=(u, v, s)$，

其中，(u, v)表示目标的位置；s 为目标尺度。为简化跟踪算法的实现，我们采用基于模板表示的颜色外观模型[38]。具体地，我们采用基于 HSV 颜色直方图外观模型，颜色直方图量化为 110 bins，其中 $N_h=N_s=N_v=10$。基于颜色外观模型的似然定义为

$$p(z_t \mid x_t) = \mathrm{e}^{-\lambda_c d^2(H^*, H(x_t))}, \tag{3.24}$$

式中，H^* 表示目标的参考模型；$H(x_t)$为候选状态 x_t 处的目标模型；λ_c 为预置参数；d 为基于颜色直方图的 Bhattacharyya 距离：

$$d(H^*, H(x_t)) = \sqrt{1-\rho(H^*, H(x_t))} \tag{3.25}$$

式中，ρ 为 Bhattacharyya 相似系数。考虑一个分块的目标表示模型(如图 3-4 所示)，目标区域 $R(x_t)$被分解为 r 个子区域，即

$$R(x_t) = \sum_{i=1}^{r} R_j(x_t)$$



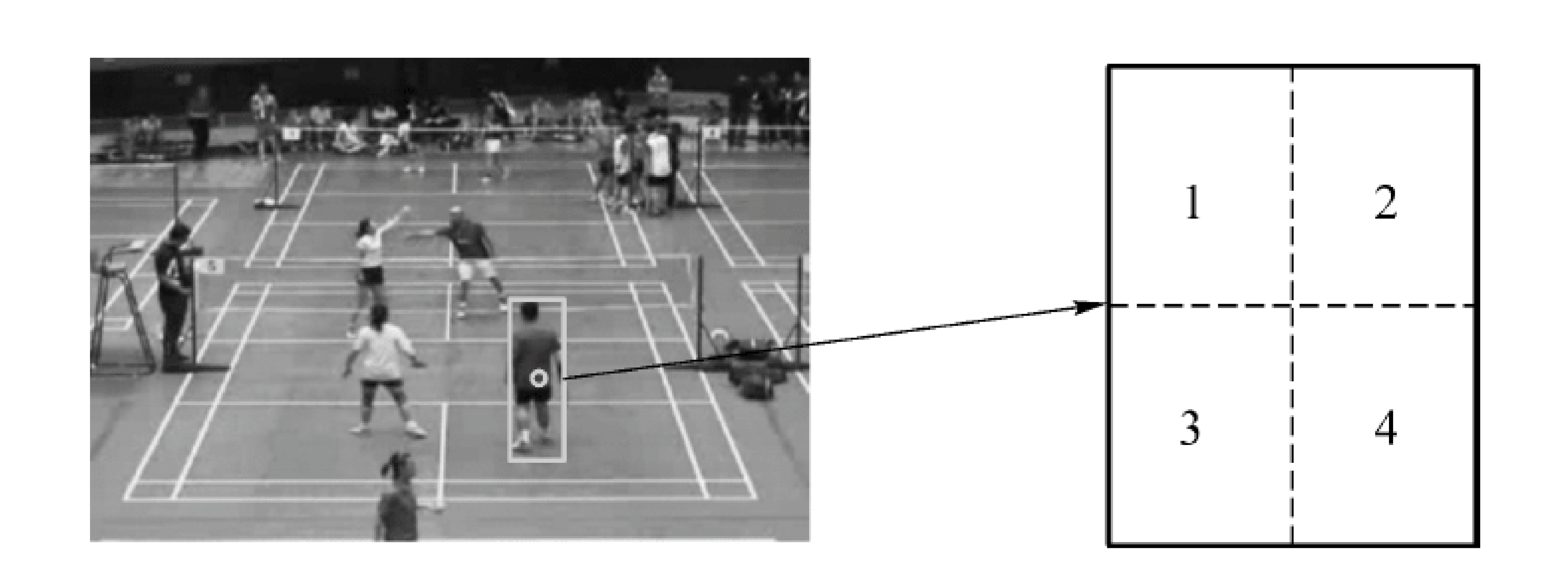


图 3-4　基于分块的目标表示。图中目标被分成四个矩形块，各分块的颜色直方图独立构建，目标观测似然由式(3.25)度量。

在这个基于分块的目标表示下，各个子区域的颜色直方图独立地构建函数。此时，外观似然可以重写为

$$p(z_t \mid x_t) = \mathrm{e}^{-\lambda_c \sum_{j=1}^{r} d^2(H^*, H(x_t))} \tag{3.26}$$

本章提出的 S-SAMC 跟踪算法能够有效地处理目标状态的不确定性，为了验证算法的这一性能，我们在算法实现中仅采用一个简单的目标运动模型作为目标状态演化的先验模型：

$$p(x_t | x_{t-1}) \sim \mathcal{N}(x_t; \mu, \Sigma) \tag{3.27}$$

式中，$\mathcal{N}(\mu, \Sigma)$为均值 $\mu = a_1 x_{t-1} + a_2 x_{t-1}$，方差为$\Sigma$的高斯分布。$a_1$，$a_2$ 定义一个常加速度 AR2 模型。在本章的实验部分可以看到，即使使用这样一个弱的运动模型，我们提出的跟踪算法也能够有效地处理目标运动的极大不确定性。

此外，S-SAMC 算法涉及多个重要参数，我们做个简要的分析。对于样本空间的划分，为简化处理通常将样本空间划分为等能量带的多个子区域。Liang 等人[93]指出，每个子区域的能量差(带宽)不应超过 2.0，即 $\Delta u_i \leqslant 2.0$，$i = 1, 2, \cdots, m-1$，以确保马氏链在各个子区域有一定的接受率。对于在各个能量子区域的期望采样频率(分布)ϕ，我们简单地设置为均匀分布，即 $\phi_i = \frac{1}{m}$，$i = 1, 2, \cdots, m$。原则上，将采样偏置到低能量区域通常可以集中计算资源在高后验密度区域，从而有助于提高整体采样效率。然而由于我们希望获得目标状态期望的无偏估计，因而将 ϕ 设置为均匀分布，详见文献[145]有关 SAMC 渐进收敛的讨论。对于 SAP 的增益因子(学习率)γ_k，式(3.16)中的 k_0 越大，通常意味着采样能够较快地访问到各个能量子区域。

3.4 实验结果

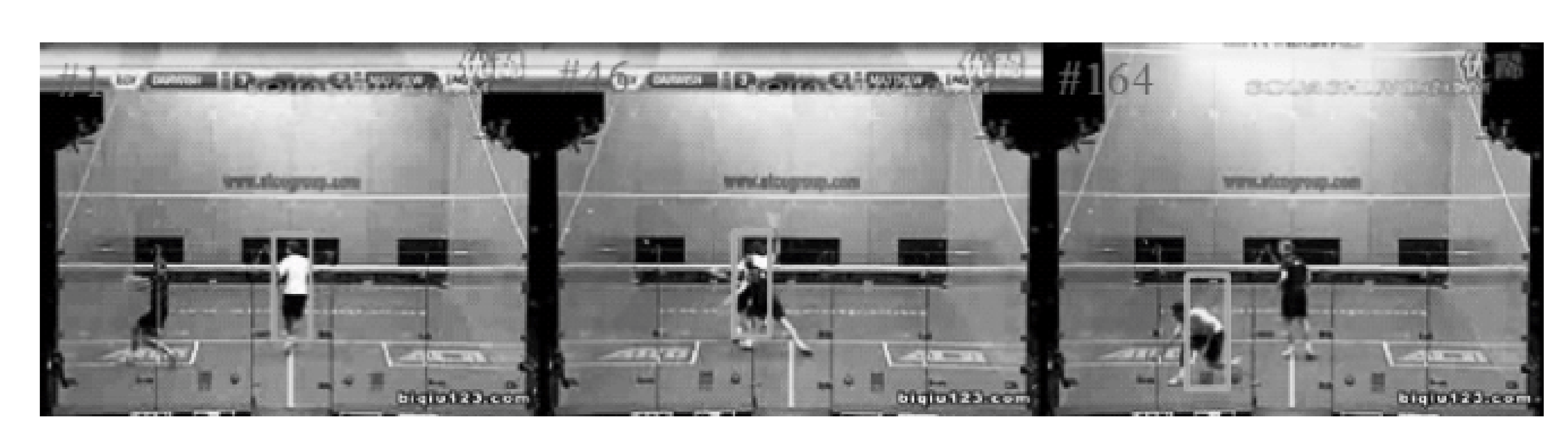
为评估提出的 S-SAMC 算法在视觉跟踪应用中的性能，在实验中使用了多个测试视频，视频场景涉及多个相似目标、多种目标运动的不确定性，包括目标运动动力学突然改变、快速运动等。在实验中，我们与几种当前主流的视觉跟踪算法进行了比较，包括 Bootstrap 粒子滤波（以下简称 BPF）[71]，基于 Mean Shift 的跟踪算法（以下简称 MS）[36]，以及基于 MCMC 采样的粒子滤波算法（以下简称 MCMC-PF）[5]。为公平比较，所有四个跟踪算法都使用相同的观测模型和似然度量。BPF、MCMC-PF 和 S-SAMC 使用相同的运动先验模型。MCMC-PF 和 S-SAMC 使用相同的提议分布。

在我们的算法实现中，似然函数的参数 λ_c 设为 20，样本空间划分为 50 个等能量带宽的子区域，即 $\Delta u_i=0.4, i=1,2,\cdots,49$。增益因子中 k_0 值设为所用的样本数量的两倍，即 $k_0=2N$。先验分布和提议分布的协方差矩阵设定为 $\Sigma=\mathrm{diag}(4^2, 4^2, 0.01^2)$。

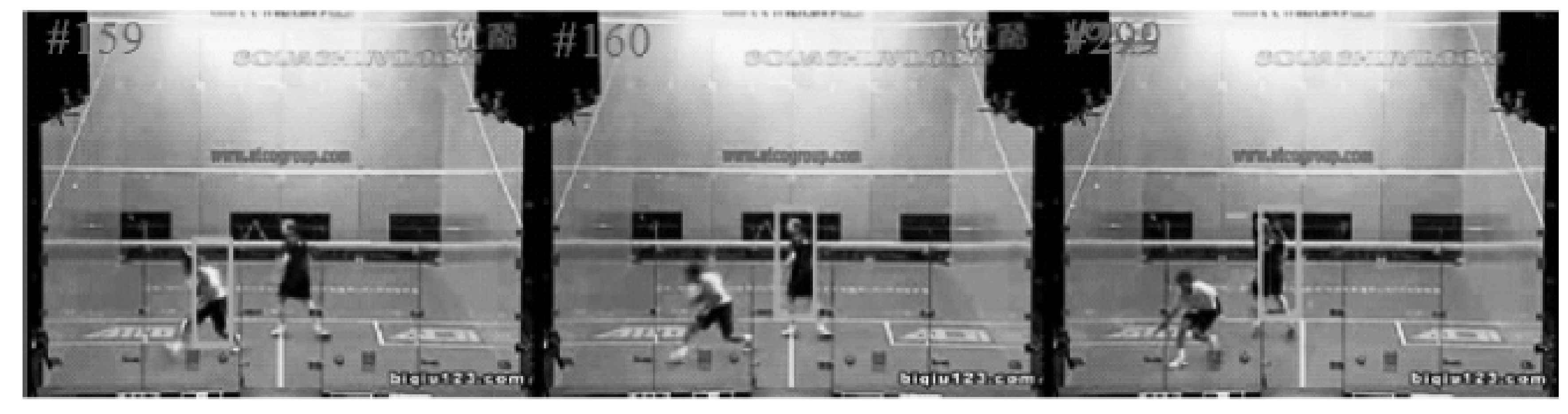
首先，我们在一个壁球比赛的视频序列 Squash 上测试四种跟踪算法的性能。Squash 视频中待跟踪的目标为穿白色上衣的运动员，它在序列中来回快速跑动击球使得其自身动力学特征具有很大的不确定性，因而用一个弱的运动模型很难有效预测和覆盖其运动范围。在这个实验中，三种基于采样的跟踪算法：BPF、MCMC-PF 和 S-SAMC 各使用 500 个粒子。图 3-5 所示为四种不同跟踪算法在 Squash 序列上的部分跟踪结果。从图 3-5 中可以看到，我们提出的 S-SAMC 算法能够有效地跟踪到快速运动的目标，当目标动力学改变时（例如突然启动或变向）S-SAMC 也能捕获这种具有很大不确定性的目标运动。同时可以看到，当两个运动目标靠近后（#35），MS 算法最先产生目标漂移，最终因陷入局部模式而导致跟踪失败，并且很难从失败中恢复。另一方面，BPF 和 MCMC-PF 算法在目标产生相互靠近、遮挡或快速运动、变向运动时容易产生跟踪漂移并导致跟踪失败，但与 MS 算法不同的是，BPF 和 MCMC-PF 能够较快地从跟踪失败中恢复目标轨迹。

其次，我们定性地评估了四种算法在另一个测试视频 Pingpong 上的性能。测试视频中乒乓球运动具有极大的不确定性。当乒乓球在下落撞击到球拍后反弹，这一碰撞前后目标的动力学特征发生突然改变。从类似的训练视频中学习一个或多个较为精确的运动模型是可能的，例如文献[86,87]，但这种方式依赖于一个离





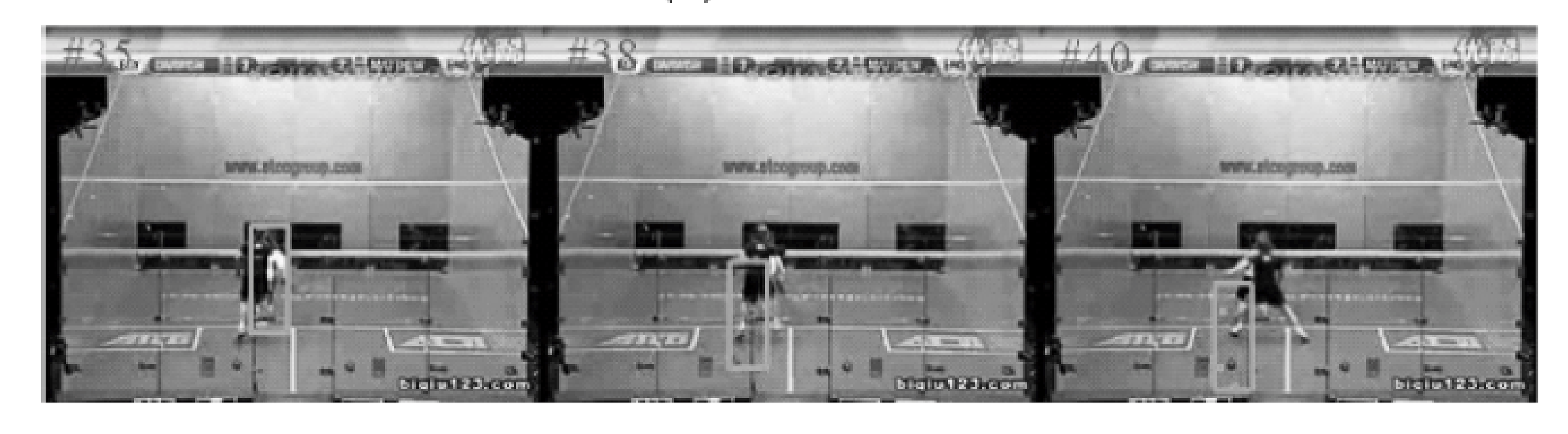
(a)S-SAMC



(b)BPF



(c)MCMC-PF



(d)MS

图 3-5　四种不同算法在视频 Squash 上的部分跟踪结果。其中，S-SAMC，BPF，MCMC-PF 各使用 500 个粒子，且使用相同的运动先验模型；四种算法使用相同的观测度量模型。

线的训练过程，因而可扩展性较差。我们的实验表明，仅仅采用一个弱运动模型，S-SAMC 算法使用 500 个粒子就可以有效地跟踪具有极大不确定性运动特征的乒乓球。部分跟踪结果如图 3-6 所示。采用相同的运动模型和外观度量，BPF 和 MCMC-PF 即使使用 1 000 个粒子，仍然无法有效地跟踪目标运动的不确定性，从而频繁地发生跟踪失败。

此外，我们在多个测试视频上定量评估了四种算法的跟踪性能。图 3-7 和图 3-8 所示为不同算法在 Squash 和 Pingpong 序列上的跟踪位置错误（像素）。表 3-1 所示为四种算法在两个测试视频上的相对位置错误（Relative Position Error，RPE）。RPE[54]定义为

$$d=\frac{\|(u,v)-(u_g,v_g)\|}{s_g},\tag{3.28}$$

式中，u_g，v_g，s_g 为测试视频中目标的真实（基准）状态，它通常通过人为标注生成。



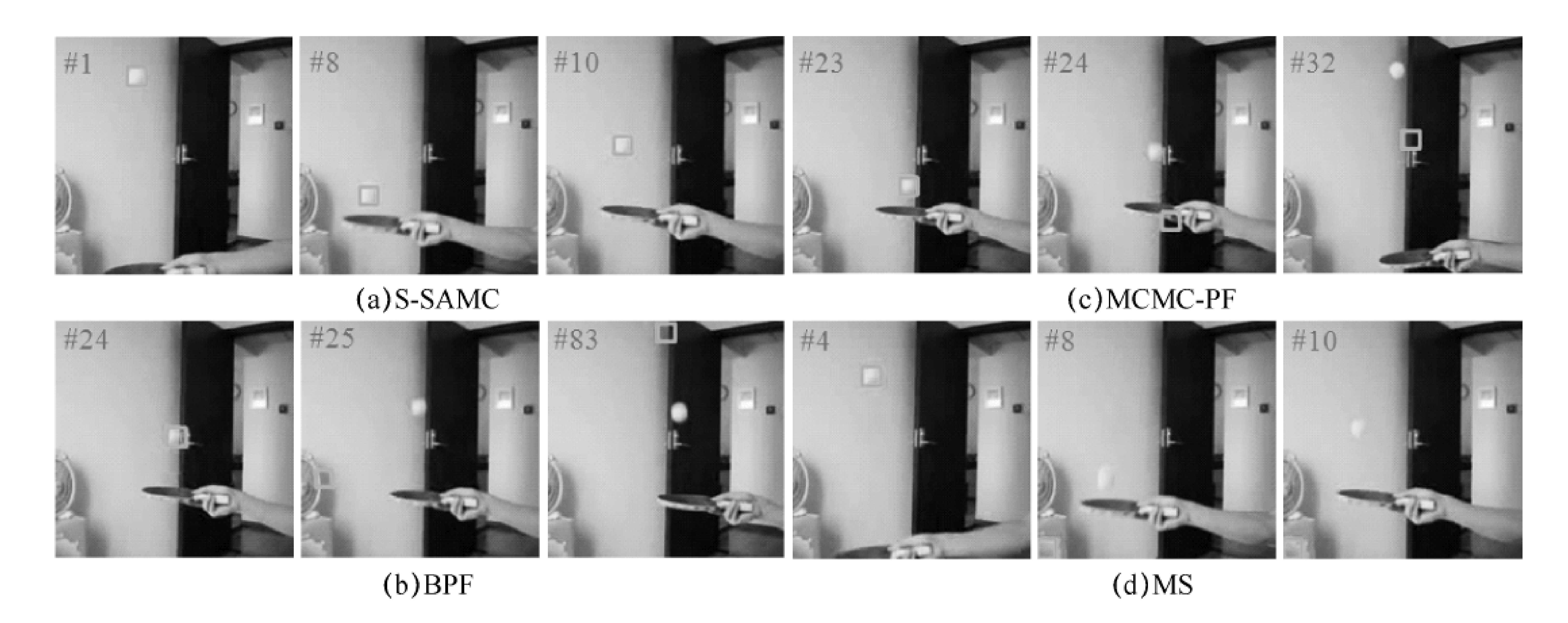


图 3-6　四种不同算法在视频 Pingpong 上的部分跟踪结果。其中，S-SAMC 使用 500 个粒子，BPF 和 MCMC-PF 各使用 1 000 个粒子；四种算法使用相同的观测度量模型。

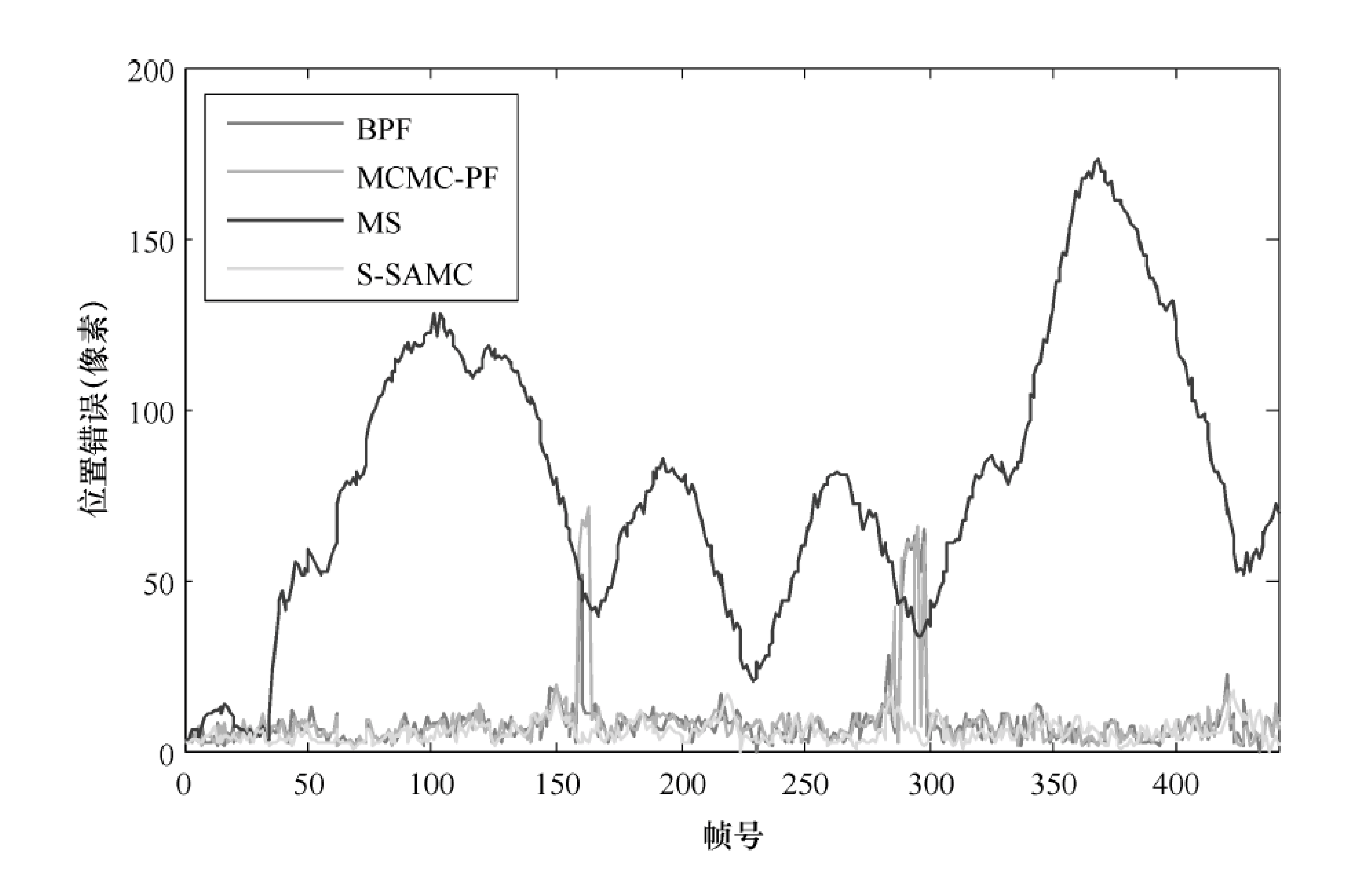


图 3-7　四种不同算法在视频 Squash 上的部分跟踪结果曲线示意。其中，S-SAMC，BPF 和 MCMC-PF 各使用 500 个粒子；四种算法使用相同的观测度量模型。

使用 RPE 度量的目的是便于在目标大小不同的情况下进行跟踪精度的比较。从图 3-7 和图 3-8 以及表 3-1 可以清楚地看到，三种基于随机采样的跟踪算法相比 MS 跟踪算法有着更好的跟踪成功率和精确度。同时，我们提出的 S-SAMC 算法相比其他三种算法在跟踪成功率、精确度以及稳定性方面都有显著的提高。



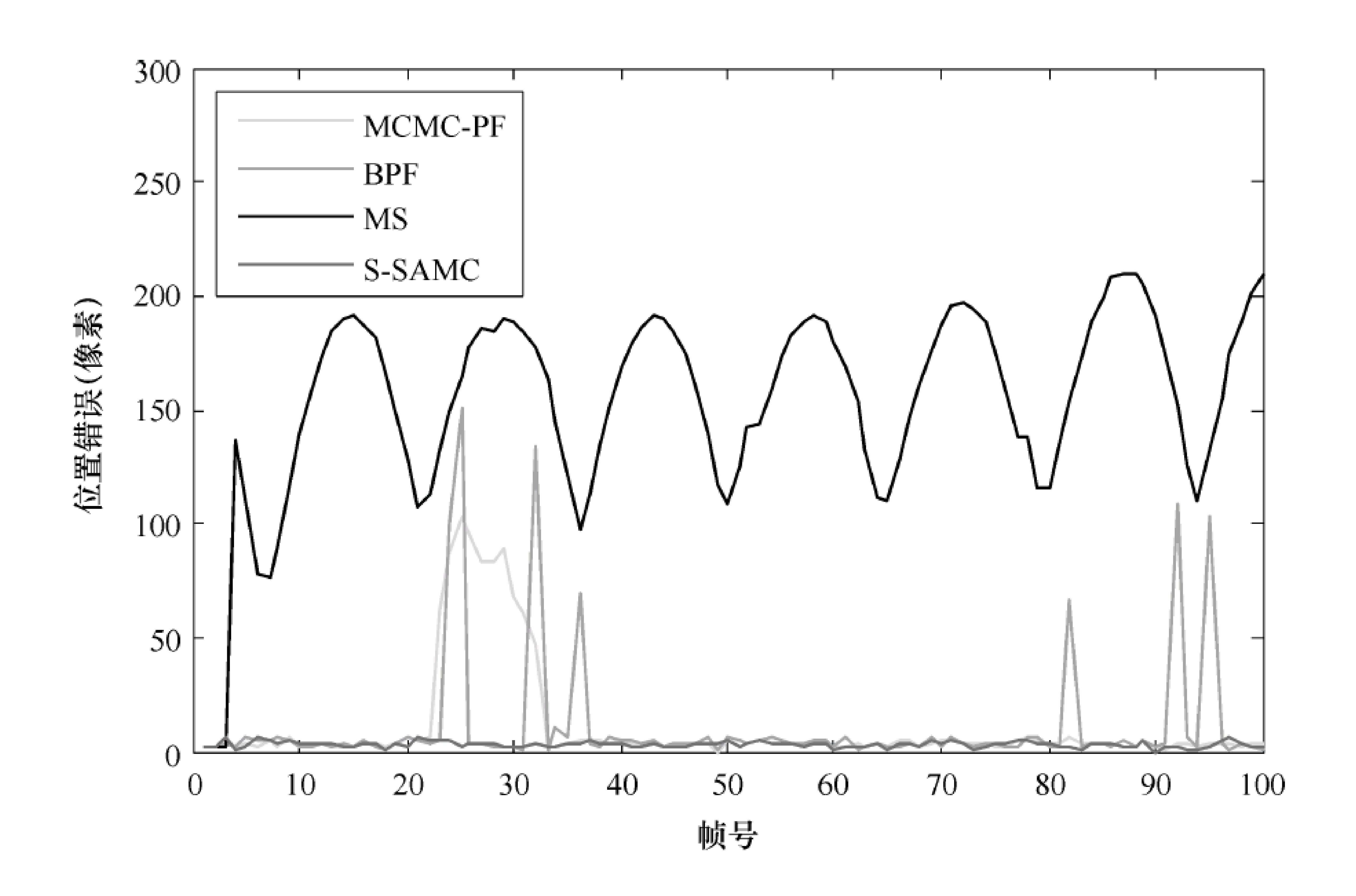


图 3-8　四种不同算法在视频 Pingpong 上的部分跟踪结果曲线示意。
其中，S-SAMC 使用 500 个粒子，BPF 和 MCMC-PF 各使用 1 000 个粒子；
四种算法使用相同的观测度量模型。

表 3-1　四种不同跟踪算法在测试视频上的跟踪性能(相对位置错误)

跟踪算法	Pingpong		Squash	
	d_μ	d_σ	d_μ	d_σ
MS	9.6386	2.6260	3.0289	1.7353
BPF	0.6515	1.6989	0.2679	0.3838
MCMC-PF	0.6619	1.4570	0.2414	0.3275
S-SAMC	0.1233	0.0735	0.1605	0.1515

3.5　小　结

本章探讨了贝叶斯滤波跟踪问题，研究了贝叶斯滤波问题的多种实现算法，包括粒子滤波及其多种变体、基于 MCMC 采样的粒子滤波算法等。通过理论分析和实验验证我们可以看到，传统的粒子滤波算法和基于 MCMC 采样的粒子滤波算法在对复杂多模的滤波分布进行序贯采样时都存在易于陷入局部模式的缺陷，这使

得采样算法往往需要使用大量的粒子近似滤波分布。本章通过将 SAMC 采样方法引入到贝叶斯滤波框架,提出一种动态加权的重要性采样算法,即 S-SAMC 采样算法,它能够有效地避免采样过陷入局部模式,从而显著地提供采样效率。基于 S-SAMC 采样的跟踪算法是一个通用的视觉跟踪框架,它易于集成更为有效地提议分布、观测模型和目标表示来进一步提高跟踪性能。

第 4 章

运动突变的视觉跟踪方法

4.1 引　　言

在视觉跟踪研究领域，大多数跟踪算法都基于一个假设，即目标运动（状态变化）和外观变化的连续性。然而，在许多复杂的场景中，这种连续性假设并不符合实际情况。如图 4-1 所示，一些体育竞技比赛中经常出现目标运动极大不确定性的情形。事实上，运动连续性假设对视觉跟踪至关重要，这是由于连续性假设通常有利于跟踪算法可以在较小的状态空间中建模目标的运动，使得在有限的计算资源条件下实现可靠的跟踪任务。一旦这种连续性假设不再成立，那么就需要在状态空间中扩大搜索（采样）范围来估计目标的状态，这通常需要增加大量的样本（粒子或迭代）来实现在庞大的方案空间中进行有效地采样。



#245



#246

图 4-1　拳击比赛视频中因镜头切换引起的目标运动突变

在视觉跟踪领域，目标的运动突变主要包括：目标自身动力学特性的极大不确定性、外在因素影响的动力学改变、摄像机镜头切换、低帧率视频等。回顾本书第 3 章提出的基于S-SAMC 采样的视觉跟踪方法。它是一个通用的视觉跟踪框架，

通过使用一个简单的运动模型和外观模型，它能够有效地处理一定程度的目标运动不确定性，例如在 Squash 和 Pingpong 视频上的实验证实了这一点。值得注意的是，S-SAMC 跟踪算法在处理不确定性动力学特征的目标跟踪问题时仍然使用了相当数量的粒子数目，这也意味着计算代价仍然较高，对于一些实时跟踪应用来说采样效率仍然有待进一步提高。为了有效地处理范围更大的运动不确定性问题，即多种运动突变的视觉跟踪问题，本章在 S-SAMC 跟踪算法的基础上研究如何进一步提高采样效率的方法和技术，实现一个能够有效处理运动突变的视觉跟踪问题的算法。我们解决这一问题的基本思想是对第 3 章提出的 S-SAMC 采样算法进一步的“自适应”，并提出一个强自适应的 MCMC 采样算法实现对滤波分布更高效地采样，从而进一步提高跟踪性能，降低计算复杂度。

4.2　相关工作

较早对目标跟踪中的运动动力学特性进行研究的是 Blake 等人[86,87]，他们从特定场景的训练数据中学习动力学特征并用于动力学变化的状态建模。这种基于运动特征学习的方法对于具有一定运动模式的特定场景下的跟踪被认为是可行的。然而由于基于特定的训练集，这种方法很难适合于一般性目标的跟踪，因而可扩展性较差，难以应用于低帧率视频等复杂运动条件下的视觉跟踪任务。多尺度和分层采样的策略是处理目标运动不连续问题的另外一种方案。在贝叶斯框架下，Sullivan 等人[88]结合不同尺度下的图像观测，提出在状态空间中进行有效搜索的跟踪方案。这种方法的缺点是，大尺度下的不准确的推断将导致在原始尺度下目标搜索的失败。针对这一困难，Hua 等人[13]提出在多尺度下的一种基于动态马尔科夫网络的合作搜索策略，并设计了一个序贯信念传播算法以实现更为准确的贝叶斯推断。

针对低帧率情况下的视觉跟踪问题，Li 等人[89]提出使用一个级联的粒子滤波器结合多个在线（或离线）学习到的检测模型来提高粒子采样效率和目标跟踪准确度。该方法在多个测试实例中取得较好的跟踪效果。然而，它需要一个事先训练好的可靠的观测模型和一个离线的学习过程，此外，它不能处理大范围快速运动的目标跟踪，如多摄像机镜头切换等。最近，Kwon 和 Lee[90]提出了一种基于 Wang-Landau 算法[91]的能够处理多种运动突变的目标跟踪算法。其基本思想是使用 Wang-Landau 算法在整个样本空间中对滤波分布进行全局采样从而避免陷入局部最优模式。由于能有效地避免在采样过程中陷入局部最优状态，这种方法在多



种不连续运动场景中取得了很好的跟踪结果。然而，该方法所使用的提议分布是基于均匀分布的全局跳跃模式，没有考虑局部模式的细节探索，尽管作者使用了一个退火模式来分阶段压缩样本空间，整个采样过程仍然需要相当数目的粒子来近似复杂的滤波分布，因而在实际应用中，离实时跟踪需要还有一定的距离。此外，尽管 Wang-Landau 算法在统计物理学中的应用取得了很大的成功，但是至今还没有坚实的理论证明 Wang-Landau 算法在一般状态空间中的收敛性，并且在一些统计学和物理学应用中已发现它只能取得有限的统计精度[93]。

4.3　方法介绍

从前面的介绍可以看到，在众多视觉跟踪算法中专门对运动突变的目标跟踪问题进行研究的工作较少，能够有效处理运动突变的视觉跟踪算法就更为稀少。实践证明[90,94]，基于自适应 MCMC 采样的视觉跟踪方法是解决这一问题的有效途径之一。本节介绍如何通过“目标分布自适应”和“提议分布自适应”构建一个强自适应的 MCMC（Intensively Adaptive MCMC，IA-MCMC）采样方法，并基于此采样策略设计一个能够有效解决运动突变的视觉跟踪算法。

4.3.1　基本思想

对于运动突变的视觉跟踪问题，场景中目标状态在前后相继两帧之间的上下文信息通常显得很微弱，或者是完全丢失了上下文信息，这使得目标跟踪变得非常困难。回想第 3 章介绍的 S-SAMC 跟踪算法，它在处理目标运动不确定性时采用了一个多元高斯作为提议分布。该提议分布是一种较为常用的处理方法，通过增大方差（即扩大延展度）可以扩大在样本空间中的采样范围，但是这种方式需要大量的粒子（迭代）以便提议分布产生的候选样本能够有更多的机会“命中”庞大样本空间中的目标模式。

有别于 S-SAMC 算法中采用的基于局部高斯扰动的提议分布，我们试图考虑一种更为有效的方案。它应该既能产生局部范围的候选样本，也能产生大范围全局跳跃的候选样本。具体到采样算法，它应该既能在局部范围内进行模式细节的提炼，也能在样本空间中多个模式之间全局跳跃，从而提高采样的混合性。基于这一思路，我们首先考虑使用一个网格（Grid）将样本空间划分为许多个单元（Cell）。假定一个 p 维的样本空间$\mathcal{X} \triangleq S_1 \times S_2 \times \cdots \times S_p$，$\mathcal{X} \subset \mathbb{R}^p$。对于每一维度 $i \in \{1,2,$

$\cdots,p$)的空间 S_i,网格将其划分为 n_i 个区间(Interval),即 $S_i=\{S_{i,j}\}_{j=1}^{n_i}$。因而,样本空间由 $n=\prod_i n_i$ 个单元组成,每个单元可以表示为

$$C \triangleq S_{1,j_1}\times S_{2,j_2}\times\cdots\times S_{i,j_i}\times\cdots\times S_{p,j_p},\quad i\in(1,2,\cdots,p\},j_i\in\{1,2,\cdots,n_i\}, \tag{4.1}$$

于是,若记 $C(x)\triangleq\{j_1,j_2,\cdots,j_i,\cdots j_d\}$ 为样本 $x=\{x_1,x_2,\cdots,x_i,\cdots,x_p\}$ 所属的单元,则有 $x_i\in S_{i,j_i}$。基于这一样本空间划分方式,我们考虑如下提议分布:

$$Q(x';x)=\lambda\mathcal{N}(x';x,\Sigma)+(1-\lambda)Q_u(x';x), \tag{4.2}$$

其中,$\mathcal{N}(x,\Sigma)$,表示均值为 $\mu=x$ 方差为 Σ 的高斯分布;$Q_u(x';x)$ 为定义于区域 $S'=\mathcal{X}\backslash C(x)$ 上的均匀分布,$\lambda\in[0,1]$ 控制提议操作在局部高斯扰动与全局随机跳跃之间的平衡。易于验证该提议分布满足转移核的全局性要求,即对任意 $x,x'\in\mathcal{X}$ 都有 $Q(x';x)>0$。图 4-2 所示为该提议分布在二维样本空间的一个描述。

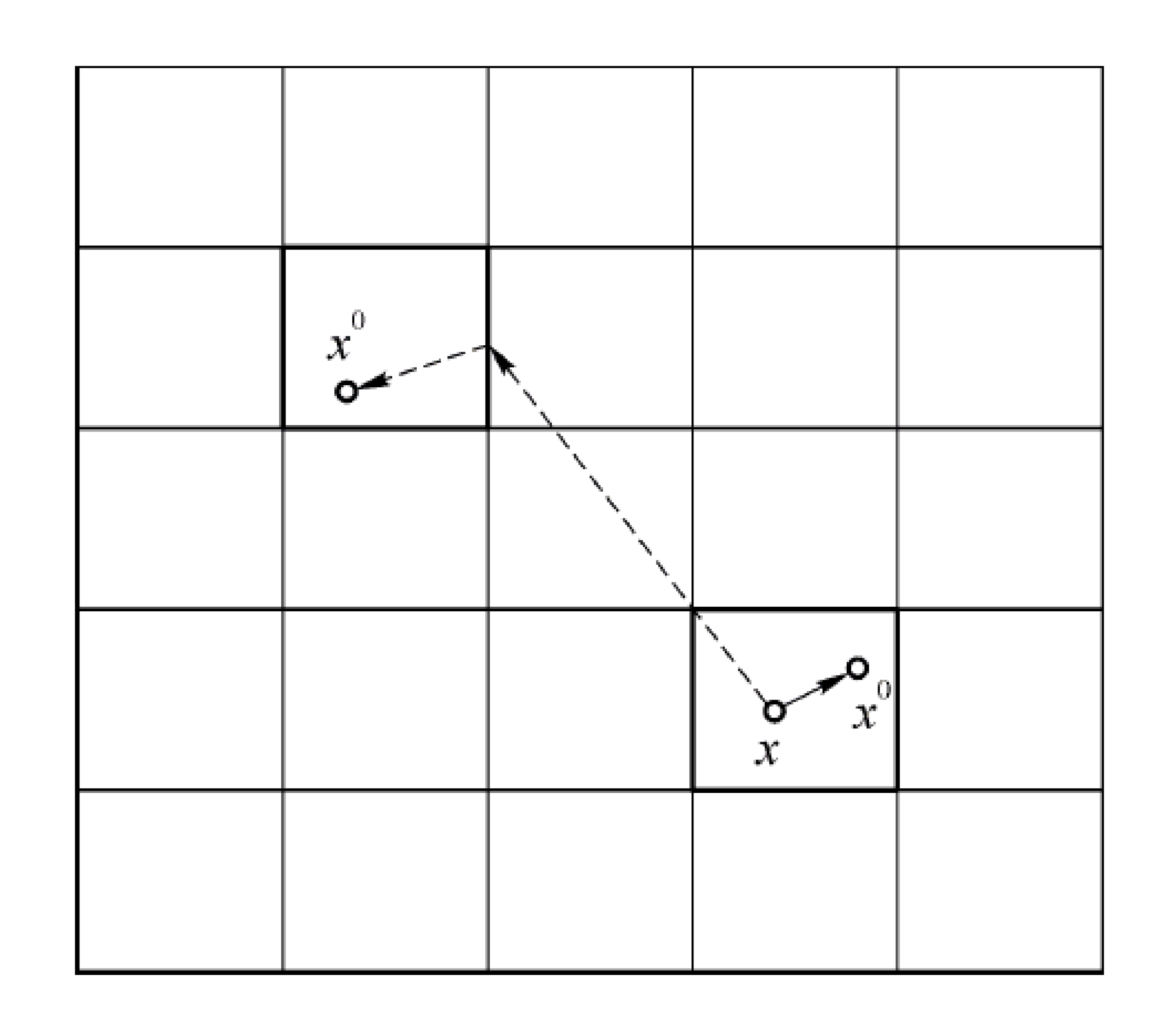


图 4-2　提议分布 $Q_u(x';x)$ 在二维样本空间的示意。样本空间被一离散网格划分为 5×5 的单元阵。该提议分布即考虑了局部随机游走(实线),也兼顾了全局随机跳跃(虚线)。

的确,在 S-SAMC 算法中使用提议分布式(4.2)替代式(3.19)能够更为有效地实现运动突变的目标跟踪(见本章实验部分)。然而,为确保跟踪的稳健性,仍然需要相当数量的粒子来覆盖运动的极大不确定性。这是由于无论是提议分布式(4.2)还是式(3.19),都采用固定的转移核。由于突变运动往往造成场景上下文信息的缺失或极其微弱,在这种情况下使用一个固定不变的转移核容易造成提议的候选样本被频繁的拒绝,尤其是在高维状态空间的情形下。为进一步提高采样效率,必须考虑更为有效的采样策略。

回顾 S-SAMC 算法在 MCMC 采样框架中使用一个 SAP 在线学习工作权重,从而有效地避免采样陷入局部模式。在此基础上,我们考虑对该采样算法进行进一步的“自适应”,具体地说,就是将其使用的不变的提议分布替换为在线学习的自适应提议分布。提议分布的学习主要基于 S-SAMC 算法构建的非齐次马氏链上的历史样本信息。直观上看,我们希望通过对历史样本信息的分析来估计样本空



间中哪些区域是“有希望的”(Promising)区域，即极有可能存在目标模式的高置信度区域。基于这样的区域划分，后续的采样过程将偏置到这些高置信度区域，从而提高采样效率。为了实现这个策略，我们在 SAMC 采样中引入一个基于密度网格的预测模型，构建一个“采样—预测—采样”的采样循环。在我们的采样策略中，密度网格携带了滤波分布在样本空间的统计信息，具体到突变运动跟踪应用中，样本空间中“有希望的”区域对应于那些高后验概率区域，即那些可能存在目标的区域。

考虑在 SAMC 采样中使用基于密度网格的预测模型的原因在于：密度网格模型力图估计样本空间的高置信度区域，使得从这些区域抽取的样本相比于从整个样本空间中均匀地抽取的样本有更大的机会靠近或达到全局最近模式点；另一方面，SAMC 采样能确保采样的全局性，避免采样陷入局部模式，并产生有代表性的样本。直观上看，这种局部预测与全局采样相结合的策略有助于提高采样的混合性。事实上，这一采样策略是 IA-MCMC 采样方法，其中目标分布自适应的任务是避免采样陷入局部模式，而提议分布自适应则力图提高采样效率。图 4-3 描述了贝叶斯滤波框架下的 IA-MCMC 采样方法的基本原理。

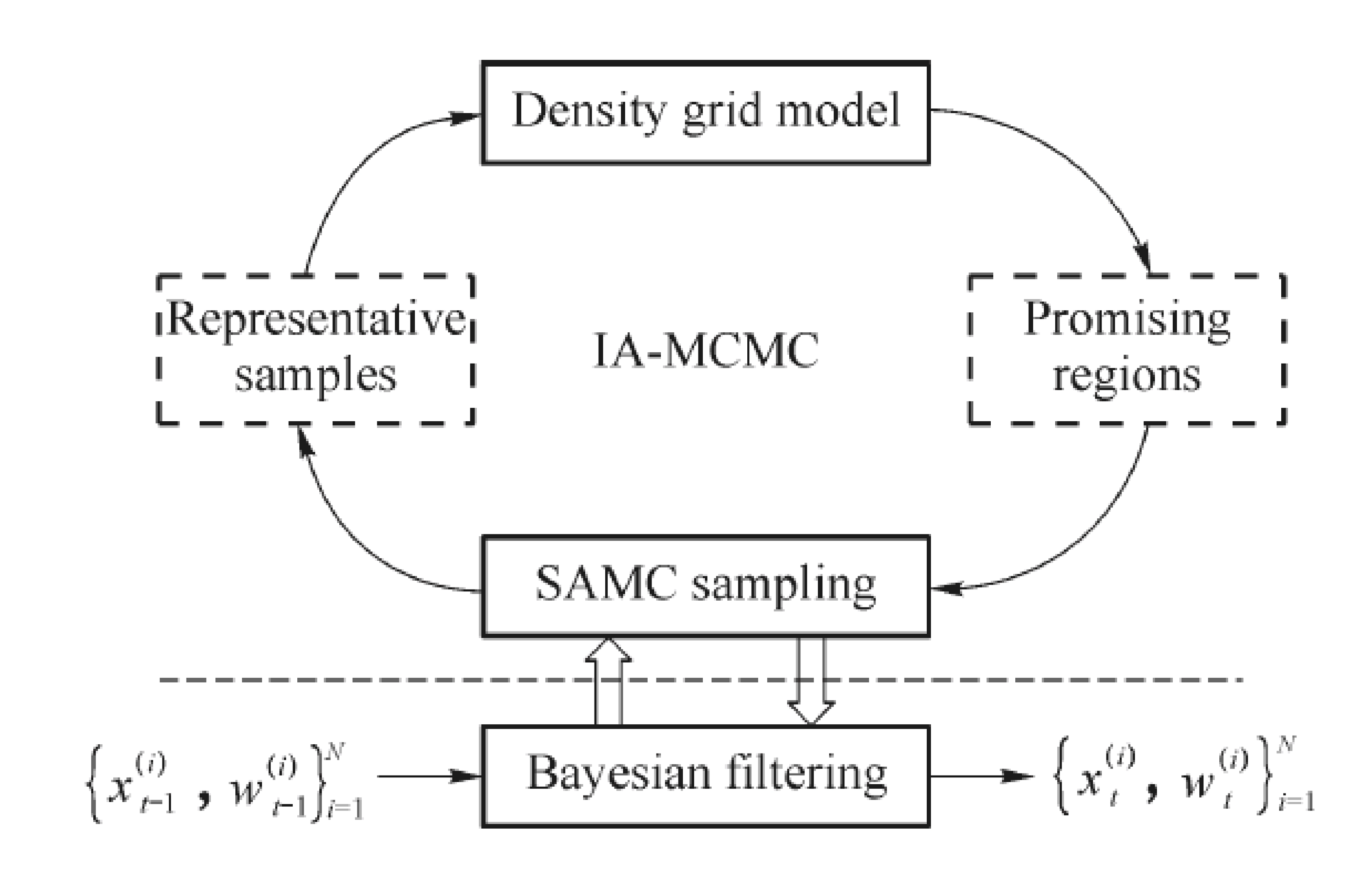


图 4-3　贝叶斯滤波框架下的深度自适应 MCMC 采样方法

4.3.2　密度网格

本节主要介绍密度网格结构及其相关操作，它是 IA-MCMC 中为实现提议分布自适应而构建的预测模型的基础。值得注意的是，本节介绍的密度网格模型与文献[146]描述的聚类模型尽管在概念上有些相似，但在模型属性和应用目的方面都有着本质的不同。后者是基于“无权重”的流数据并用于聚类，而前者则使用带有滤波分布信息(能量)的样本进行高置信度区域预测。

密度网格结构是在本章 4.3.1 节介绍的网格结构通过引入单元(Cell)的密度

概念扩展而成的。令 G 表示一个密度网格结构,其中每一个单元初始化值为 0。对于每一个由S-SAMC 算法抽取的样本 x_t,为其赋予一个密度系数,即

$$d(x_t) = \mathrm{e}^{-E(x_t)} \tag{4.3}$$

式中,$E(x_t)=-\log p(x_t|z_{1:t})$,表示滤波分布的能量函数。密度系数的概念的直观意义是,在采样过程中密度网格模型为具有较低能量值的样本赋予较大的权重,以便更有效地发现和提炼样本空间中的高置信度区域。基于密度系数的概念,密度网格中每一单元的密度值可以方便地计算得到。对于单元 C,在第 k 次迭代时其密度等于落入该单元的所有历史样本的密度系数之和,即

$$D(C,k) = \sum_{i=1}^{k} d(x^{(i)}) I(x^{(i)} \in C) \tag{4.4}$$

现在考虑基于此密度网格结构构建一个预测模型。我们定义具有较高密度值的单元为高置信度区域。令 $D_0^{(k)}$ 表示在第 k 次迭代时的密度阈值,此时高置信度区域定义为

$$\Psi^{(k)} = \{C: D(C,k) \geqslant D_0^{(k)}\} \tag{4.5}$$

典型地,我们将密度阈值 $\mathrm{D}_0^{(k)}$ 设定为第 k 次迭代时所有单元密度值的平均。随着采样的进行,样本空间中高置信度区域会连续地得到更新和提炼,这是因为在采样过程中每个单元的密度都是动态变化的,密度阈值也是动态更新的。从数据挖掘的观点来看,这一密度网格模型可以看作一个在线的预测模型,它把样本空间动态地划分为高置信度区域和非高置信度区域。这一模型尽管较为简单,但对于学习自适应的提议分布却很有效。

4.3.3 提议分布自适应

如图 4-3 所示,在 IA-MCMC 采样方法中存在两个在线学习过程:工作权重(即态密度项)的估计是通过 SAMC 采样算法中的一个 SAP 实现的,它使采样避免陷入局部模式;在这一节里,我们将讨论借助密度网格模型实现另一个在线学习过程,即提议分布自适应,其目的是学习更加有效的提议分布并偏置采样到样本空间中高置信度区域,从而提高采样效率。

由密度网格预测得到的高置信度区域可能包含多个密度单元。记 $n^{(k)}$ 为高置信度密度单元的个数;$C_i^{(k)}$ 为在第 k 次迭代时第 i 个高置信单元。令 $Q_i^{(k)}$ 表示第 k 次迭代时与第 i 个高置信度区域相关联的提议分布。给定当前样本 $x_t^{(k)}$,原则上,我们希望以一个较大的概率值从高置信度区域中抽取一个样本。考虑如下的提议分布:



$$Q_i^{(k)}(\cdot;x_t^{(k)}) = \beta q_i^{(k)}(\cdot;x_t^{(k)}) + (1-\beta)q^{(k)}(\cdot;x_t^{(k)}), \quad i = 1,2,\cdots,n^{(k)}, \tag{4.6}$$

其中,β 为从高置信度区域抽取样本的概率;$q_i^{(k)}$ 表示直接从第 i 个高置信度区域产生样本的提议分布;$q^{(k)}$ 表示从非高置信度区域采样的提议分布。基于这个定义,用于 SAMC 采样的提议分布定义为

$$Q^{*}(\cdot;x_t^{(k)}) = \sum_{i=1}^{n^{(k)}} \frac{1}{n^{(k)}} Q_i^{(k)}(\cdot;x_t^{(k)}). \tag{4.7}$$

这一提议分布兼顾了局部随机游走和全局的随机跳跃。同时,它偏置采样到高置信度区域来提高采样效率。值得注意的是,这个提议分布也能在非高置信度区域产生有限数量的候选样本,帮助采样避免陷入局部模式。

由式(4.7)可以看到,提议分布是随着采样不断变化的。在采样开始阶段,由于高置信度区域尚未被挖掘出来,因而我们以一个较大的概率值学习和更新提议分布,随着采样过程的深入,学习到的提议分布渐进地靠近“最佳”提议分布,因而对提议分布的更新就不必像初始采样阶段那么频繁。我们考虑如下的提议分布更新策略:

$$Q^{(k)} = \xi_k Q^{*} + (1-\xi_k)Q^{(k-1)} \tag{4.8}$$

式中,Q^{*} 表示在第 k 次迭代时学习到的提议分布;$Q^{(k-1)}$ 表示在 $k-1$ 次迭代时使用的提议分布;ξ_k 为学习率。此处定义:

$$\xi_k = (k - M_0)^{-\eta}, k > M_0 \tag{4.9}$$

式中,η 为控制学习率的下降速度。

4.3.4　强自适应的 MCMC 采样

IA-MCMC 采样算法包含两个采样阶段,即初始化阶段和自适应采样阶段。初始采样阶段的任务是挖掘能量分布的粗略模式(Rough Mode),而自适应采样阶段则在样本空间中提炼高置信度区域并指导采样达到或靠近全局最佳模式。

在初始采样阶段,我们采用固定不变的转移核并使用一定数量的样本调用 SAMC 采样算法,以获得样本空间中能量分布的粗略模式。基于初始采样阶段获取的样本集,密度网格构建一个初始的预测模型,它把样本空间粗略地划分为高置信度区域和非高置信度区域。在自适应采样阶段,为有效地提炼和求精高置信度区域,我们将该采样阶段继续划分为多个小阶段。随着采样的进行,通过使用密度网格,高置信度区域逐渐地被压缩和提炼。特别地,我们使用一种类似退火(Annealing)地方式压缩“有希望的”区域,即一旦采样进入一个新的阶段,就重新构建

一个具有更高分辨率的密度网格，例如，可以把每个密度单元切分为等体积(面积)的小单元并重新计算单元的密度。整个 IA-MCMC 采样过程如图 4-4 所示。基于 IA-MCMC 采样的序贯贝叶斯滤波跟踪算法描述如下：

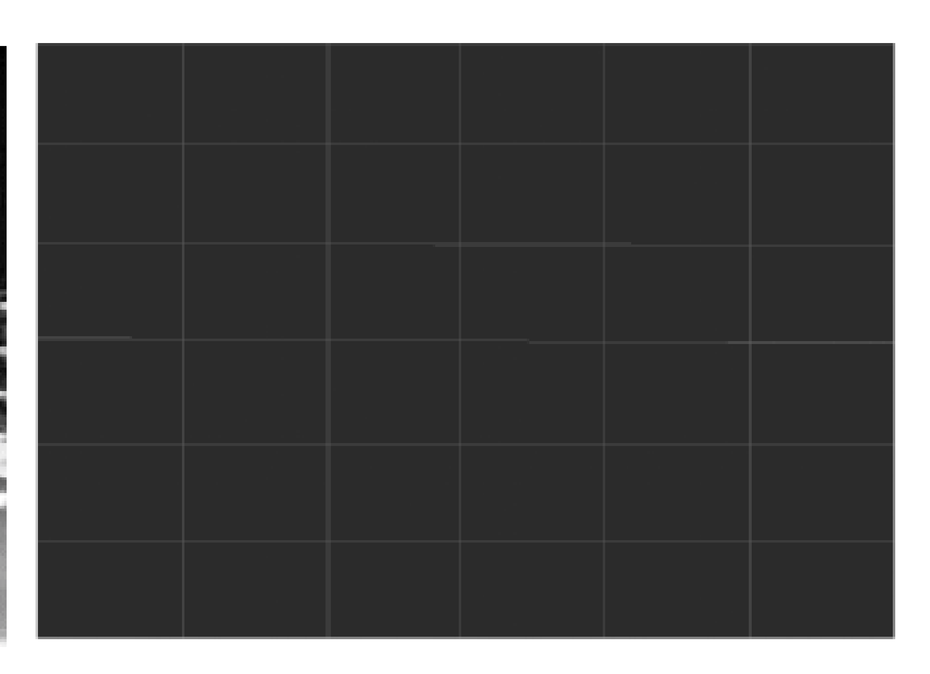
算法输入：粒子集$\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$；

算法输出：粒子集$\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^{N}$，状态估计 $\hat{x}_t$；

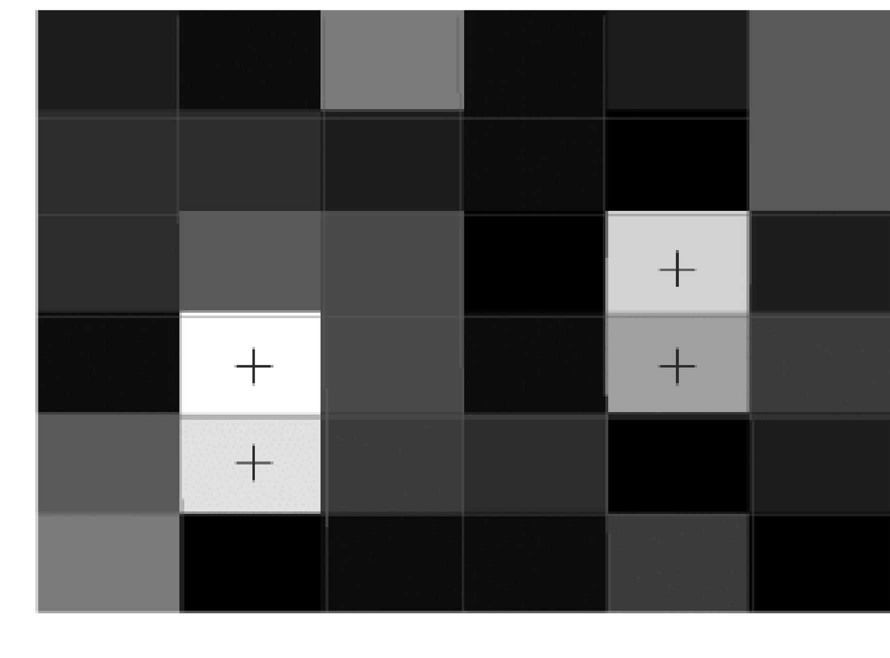




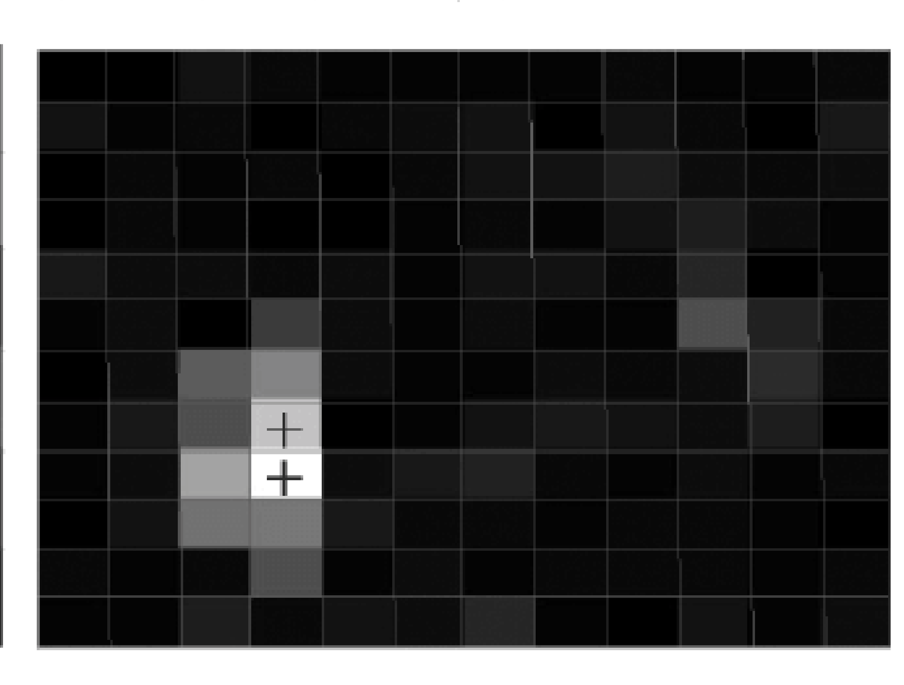
(a)因镜头切换引起的目标运动突变



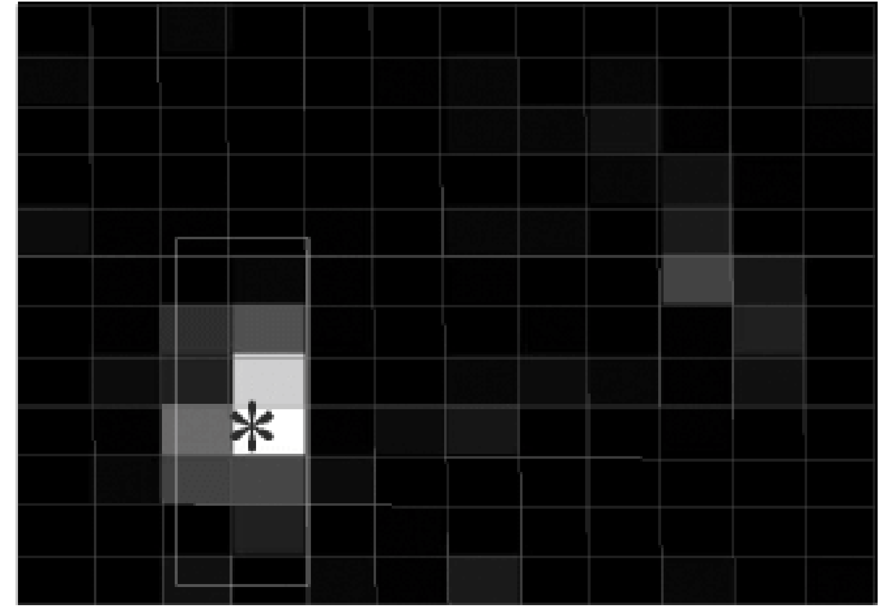
(b)初始的6×6密度网格



(c)初始采样阶段完成后粗略的“有希望的”区域（标记为“+”）



(d)随着采样的进行，“有希望的”区域不断地被压缩和提炼



(e)最后估计的目标状态

图 4-4　IA-MCMC 采样过程。

(1) 初始化

- 样本空间划分：$\mathcal{X}=\{\mathcal{X}_i\}_{i=1}^{m}$
- 初始化工作权重：$g_i=1, i=1,2,\cdots,m$
- 初始化密度网格：$s=0; k=0, D(C,k)=0, \forall C\in G$

(2) 重采样$\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^{N}$，产生等权重粒子集$\left\{x_{t-1}^{(i)}, \dfrac{1}{N}\right\}_{i=1}^{N}$；

(3) 按照(3.18)式产生初始样本 $x_t^{(0)}$；

(4) For $k=1$ to M_0　//初始采样阶段：

- 按式(4.2)抽取候选样本 x'；
- 按照式(3.20)计算接受比 α；
- 以概率 α 接受 x'，并令 $x_t^{(k)}=x'$；否则令 $x_t^{(k)}=x_t^{(k-1)}$；
- 按式(3.14)更新工作权重；

End For



(5) For $k=M_0+1$ to N　//自适应采样阶段：

- 若采样进入一新的阶段，则：
  - 使用更高的分辨率重建密度网格；
  - 更新每个单元的密度值；
  - $s=s+1$；
- 按式(4.8)抽取候选样本 x'；
- 按照式(3.20)计算接受比 α；
- 以概率 α 接受 x'，并令 $x_t^{(k)}=x'$；否则令 $x_t^{(k)}=x_t^{(k-1)}$；
- 按式(3.14)更新工作权重；
- 更新密度网格的各个单元密度值；

End For

(6) 粒子加权：

- 按式(3.21)对样本集加权；
- 归一化粒子集；

(7) 基于粒子集估计目标状态 $\hat{x}_t$。

事实上，IA-MCMC 算法是一个嵌入数据挖掘模式的自适应采样方法。我们在采样过程中使用这样一个密度网格模型辅助学习提议分布，很大一方面在于其高效和易于实现，只需极低的计算代价。由于该预测模型在自适应采样过程中被频繁地使用，因而使用一个高效的预测模型是必要的。我们也考虑了其他一些具有更强预测性能的数据挖掘模型，如常用的 CART 模型[118]。然而，在我们的视觉跟踪应用中使用这些分类与预测模型需要耗费更高的计算资源。

最后需要指出的是，我们提出的 IA－MCMC 算法理论上属于适应的 MCMC (A-MCMC)采样算法的范畴[93,102]，因而除了必须遵从 SAMC 采样算法收敛性条件外，还需满足渐减自适应条件，确保采样算法的各态历经性。按照自适应 MCMC 算法的收敛条件[93, 102]易于验证本章提出的 A-MCMC 算法的各态历经性。

4.3.5　实现细节

本节讨论 IA-MCMC 算法应用于运动突变的视觉跟踪的实现细节。为简化表述，本书第 3 章提出的 S-SAMC 跟踪算法称为 SAMC 跟踪算法，本章提出的基于 IA-MCMC 采样的跟踪算法称为 IA-MCMC 跟踪算法。与 SAMC 跟踪算法类似，目标状态表示为$x=\{u,v,s\}$，同样使用基于颜色直方图的外观模型。由于面向运动突变的视觉跟踪任务，我们定义一个更弱的运动模型：

$$p(x_t \mid x_{t-1}) = \begin{cases} U(\mathcal{X}_p) & \text{for object position} \\ \mathcal{N}(x_t^s; \mu_s, \sigma_s^2) & \text{for object scale} \end{cases} \tag{4.10}$$

式中,$U(\mathcal{X}_p)$为定义在二维空域$\mathcal{X}_p$上的均匀分布;$\mathcal{N}(\mu_s, \sigma_s^2)$为均值为$\mu_s = a_1 x_{t-1}^s + a_2 x_{t-2}^s$,方差为$\sigma_s^2$的高斯分布,其中$a_1$,$a_2$为常加速度AR2模型。在实际应用中,使用这样一个弱运动先验是合适的,由于运动突变的目标可能在场景中的任何位置出现,而目标的尺度变化则可以近似地认为是缓慢(光滑)地变化。类似地,对于提议分布中的尺度维,也可采用这种光滑性假设。

样本空间的划分方式与SAMC跟踪算法中的设置方式一致,期望采样分布也设定为均匀分布。由于自适应采样阶段已经将采样偏置到高置信度区域,因而不必要设置偏置的期望采样分布。此外,密度网格的初始分辨率设为6×6。$\lambda=0.2$,使初始采样阶段更偏向于全局随机跳跃,以便于快速地探索样本空间的能量分布的粗鲁模式。$\beta=0.9$,使提议分布基本集中在高置信度区域采样,以提高采样效率。

4.4 实验结果

为了全面评估IA-MCMC跟踪算法的性能,我们收集了十几个涉及各种突变运动场景的测试视频。未经优化,IA-MCMC跟踪算法(300个粒子)在主频为2.8 GHz的PC上获得大约20 f/ps的帧率。鉴于算法运行环境的差异,对不同跟踪算法在运行速度和计算代价方面的比较,考察算法所使用的粒子数目(样本数目或迭代次数)更为客观和公平,因为本章参与性能比较的跟踪算法的计算代价主要耗费在迭代的似然计算上。

在实验中,我们首先比较本书提出的SAMC和IA-MCMC两种跟踪算法。通过定性和定量的比较,考察提议分布自适应能够对跟踪算法的性能带来什么样的影响。然后,我们将IA-MCMC与另外三种基于MCMC采样的跟踪算法进行定性和定量地比较,这三种基于MCMC采样的跟踪算法分别是:MCMC-PF[5],A-WLMC[90]和基于AM采样[101]的跟踪算法。为公平比较,所有跟踪算法采用相同的运动先验和观测似然模型。

4.4.1 提议分布自适应的影响

我们首先测试SAMC和IA-MCMC在视频Squash上的跟踪性能。该视频是从一个超过400帧的壁球比赛视频中按每15帧抽取1帧的方式生成的低帧率视



频。如图 4-5 所示，使用 300 个粒子 IA-MCMC 能够成功地估计目标的极大不确定性运动，而使用相同粒子数的 SAMC 却在一些帧中没能有效地估计到目标的突变运动。为揭示提议分布自适应能给 SAMC 采样效率带来多大程度的提高，我们在视频 Squash 和 Pingpong 上定量地评估它们的跟踪精确度。如图 4-6 所示，使用相同数目的粒子，IA-MCMC 获得更高的跟踪精确度和成功率。这种跟踪性能的提高主要归功于 IA-MCMC 在采样过程中通过提议分布自适应使采样偏置于高置信度区域，从而提高整个采样效率。

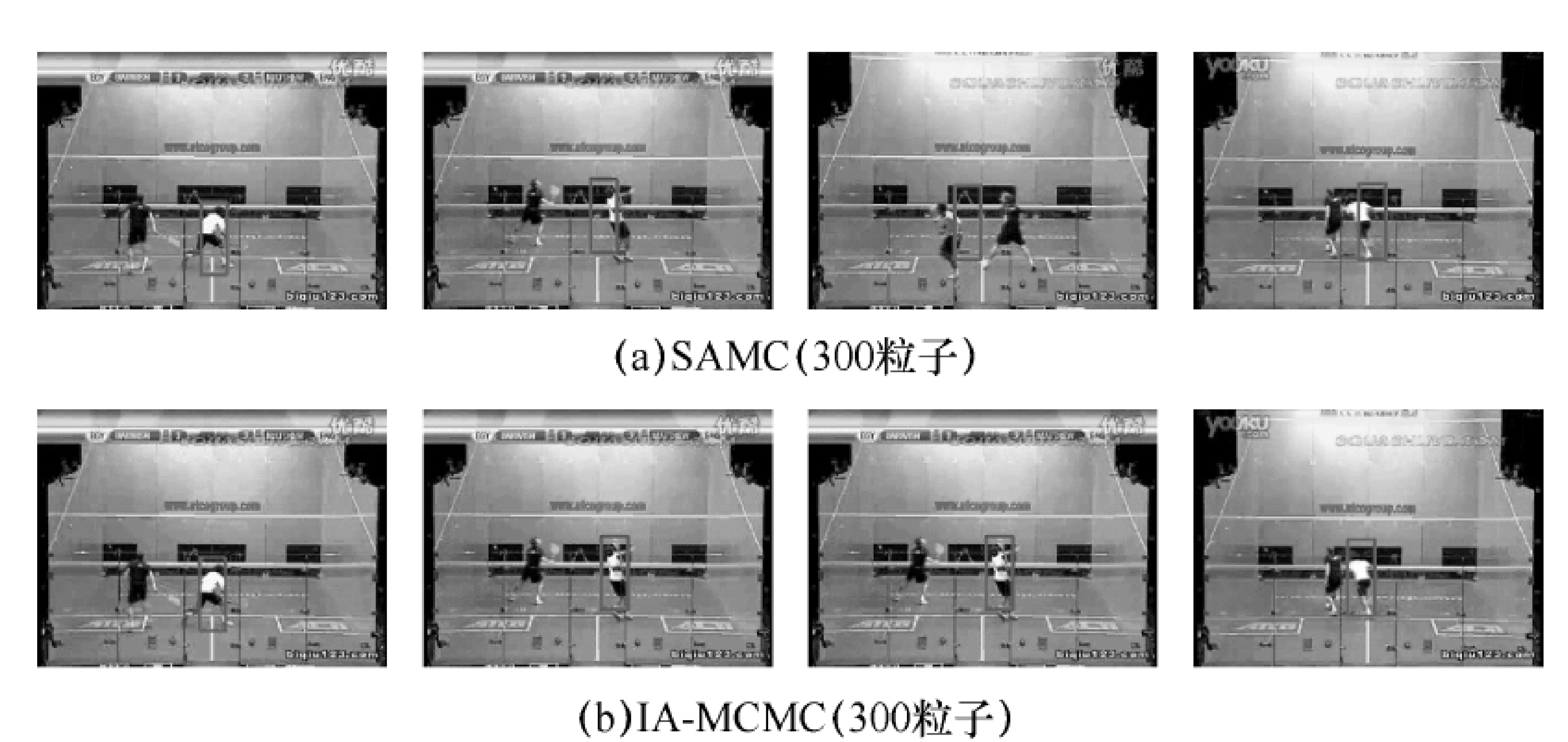


图 4-5　SAMC 和 IA-MCMC 算法在视频 Squash 上的部分跟踪结果。

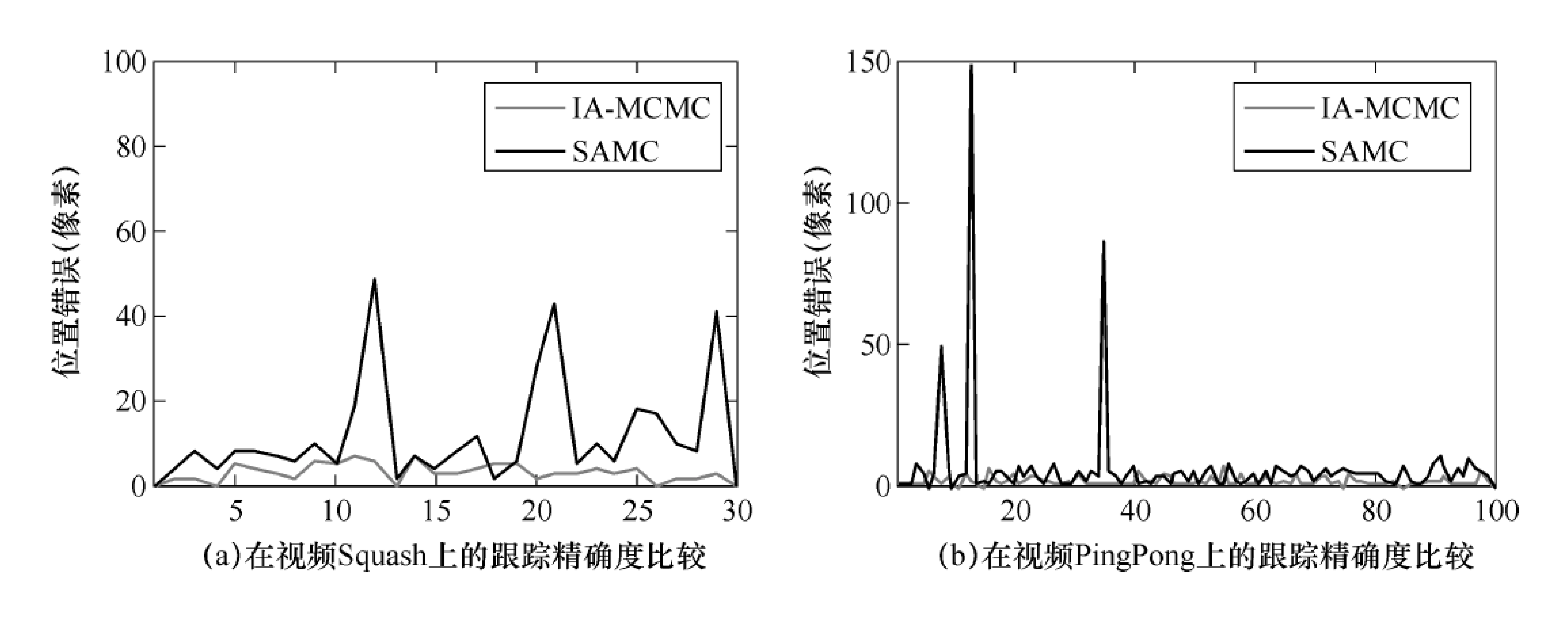


图 4-6　SAMC 和 IA-MCMC 算法在视频 Squash 和 Pingpong 上的跟踪精确度比较

4.4.2　定性比较

为评估 IA-MCMC 跟踪算法的性能，我们在涉及多种运动突变(包括目标快速移动、目标动力学突变、镜头切换以及低帧率视频)的多个视频序列上与其他三种基于 MCMC 采样的跟踪算法进行定性地比较。

1. 目标快速移动

我们首先在 Face 视频[147]上定性评估四种跟踪算法的跟踪性能。在该视频中,被跟踪的人脸快速地左右移动。MCMC-PF 和 AM 算法使用 50 个粒子,A-WLMC 和 IA-MCMC 使用 25 个粒子,部分跟踪结果如图 4-7 所示。从中我们可以看到,仅仅使用少量的粒子,A-WLMC 和 IA-MCMC 能够有效地跟踪到快速移动的目标,而 MCMC-PF 和 AM 尽管使用了 2 倍的粒子,但仍然频繁地出现跟踪失败。

2. 镜头切换

视频场景中摄像机镜头切换往往导致目标的突变运动。我们在 YoungKi 视频上定性评估四种跟踪算法的性能。在这个视频中,除了当镜头切换时目标运动连续性被破坏外,被跟踪的行人大部分时间段内运动都是近似连续的。当发生镜头切换时,目标在前后帧之间的关联信息几近丢失。实验表明,仅使用 100 个粒子,IA-MCMC 和 A-WLMC 能够有效地跟踪到因镜头切换导致的运动突变的目标,部分跟踪结果如图 4-8 所示。另一方面,使用1 000个粒子,MCMC-PF 和 AM 仍然不能有效地估计到目标的突变运动,当镜头切换发生时往往导致目标跟踪的失败。其主要原因在于 MCMC-PF 和 AM 缺乏有效的机制避免采样陷入局部模式。对于 YoungKi 视频,A-WLMC 算法能够有效地处理因镜头切换引起的运动突变,但同时我们也看到在跟踪精确度方面,使用相同的粒子数,IA-MCMC 比 A-WLMC 获得更高的跟踪精确度,如图 4-8(c)和 4-8(d)所示。

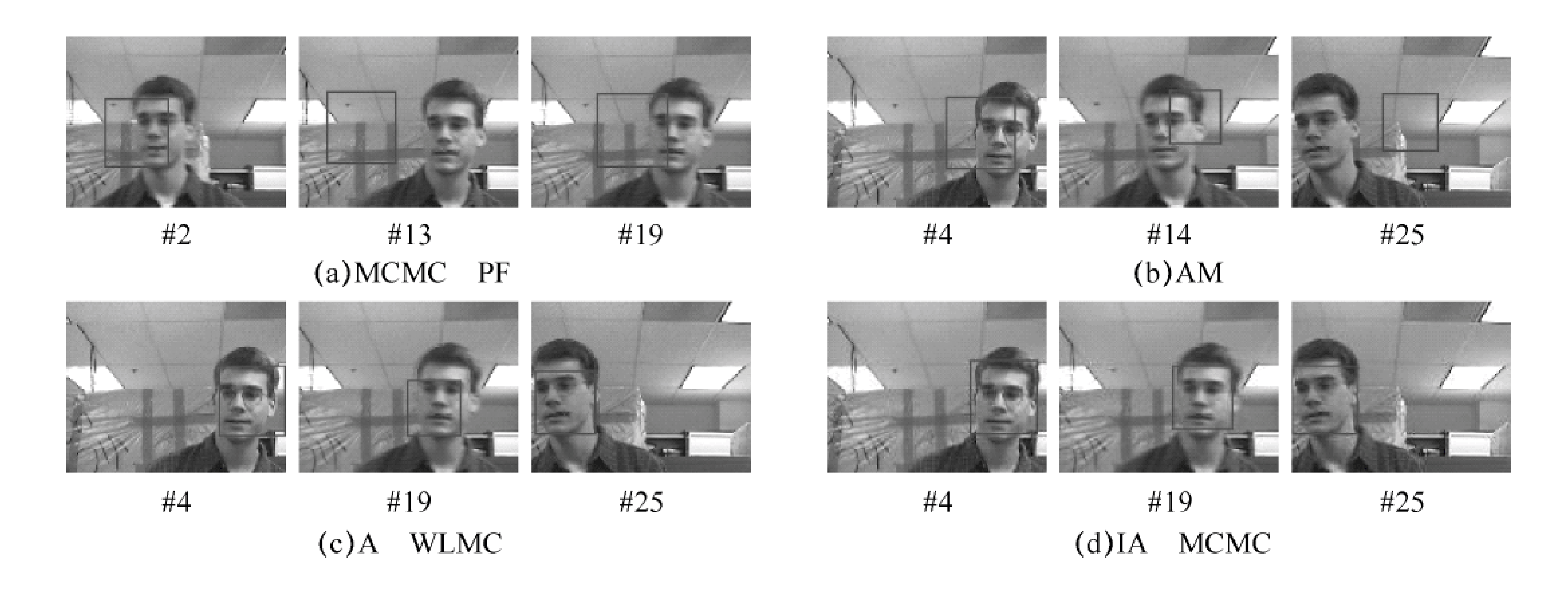


图 4-7 四种不同跟踪算法在视频 Face 上的部分跟踪结果。A-WLMC 和 IA-MCMC 使用 25 个粒子,MCMC-PF 和 AM 使用 50 个粒子。



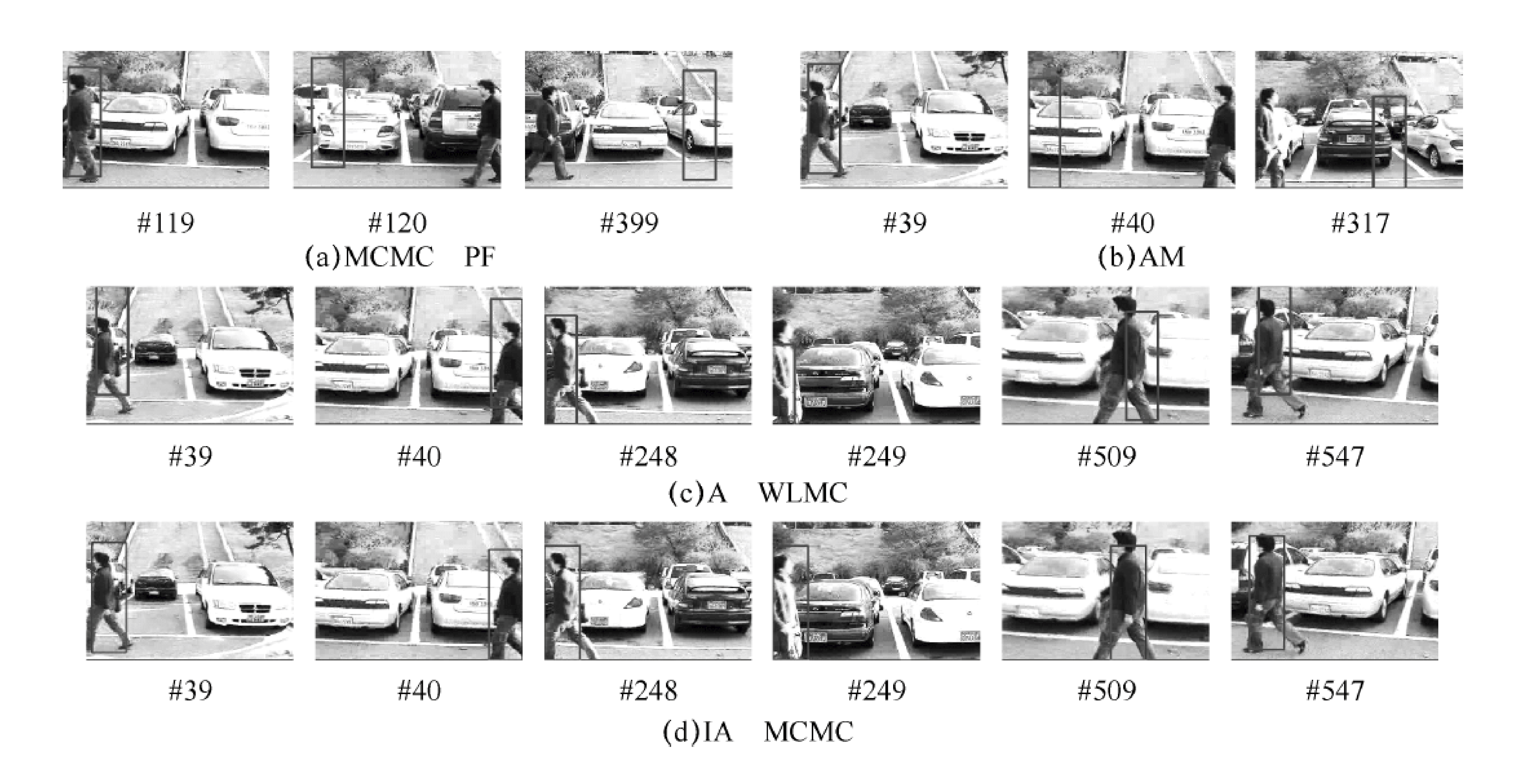


图 4-8　四种不同算法在视频 YoungKi 上的部分跟踪结果。MCMC-PF 和 AM 使用 1 000 个粒子，A-WLMC 和 IA-MCMC 使用 100 个粒子。

3. 动力学突变

我们在 Pingpong 视频上测试四种不同算法在目标动力学发生突变情况下的跟踪性能。当乒乓球在下落碰到球拍后反弹，这一碰撞前后目标的动力学特征发生突然改变。这种难以预测的目标动力学特征使跟踪任务变得相当困难。我们的实验表明，使用 300 个粒子 IA-MCMC 能够有效地在整个序列中跟踪到乒乓球的不确定性运动，部分跟踪结果如图 4-9 所示。值得注意的是，该视频场景中目标的尺寸远小于前一个实验中 YoungKi 视频的行人。因而，样本空间中“有希望的”区域相对的小得多，这意味着需要更多的粒子去挖掘“有希望的”区域以提高采样效率。另一方面，即使使用 1 000 个粒子，MCMC-PF 和 AM 算法在这个测试视频中仍然表现较差，频繁出现目标漂移和跟踪失败。此外，同样使用 300 个粒子，A-WLMC 算法在该视频的一些帧中(例如 ＃50，＃68，＃89)存在跟踪失败现象。

4. 低帧率视频

这个实验的任务是在一段低帧率的网球比赛视频 Tennis 中跟踪一个具有极大不确定性运动特征的网球选手。该视频是从一段超过 700 帧的原始比赛视频经过每 20 帧抽取 1 帧生成的低帧率视频。我们评估了四种跟踪算法在这个视频序列上的跟踪性能，部分跟踪结果如图 4-10 所示。从图 4-10 可以看到，由于严重的丢帧(Frame Dropping)，即使使用1 000 个粒子，MCMC-PF 和 AM 仍然频繁地出

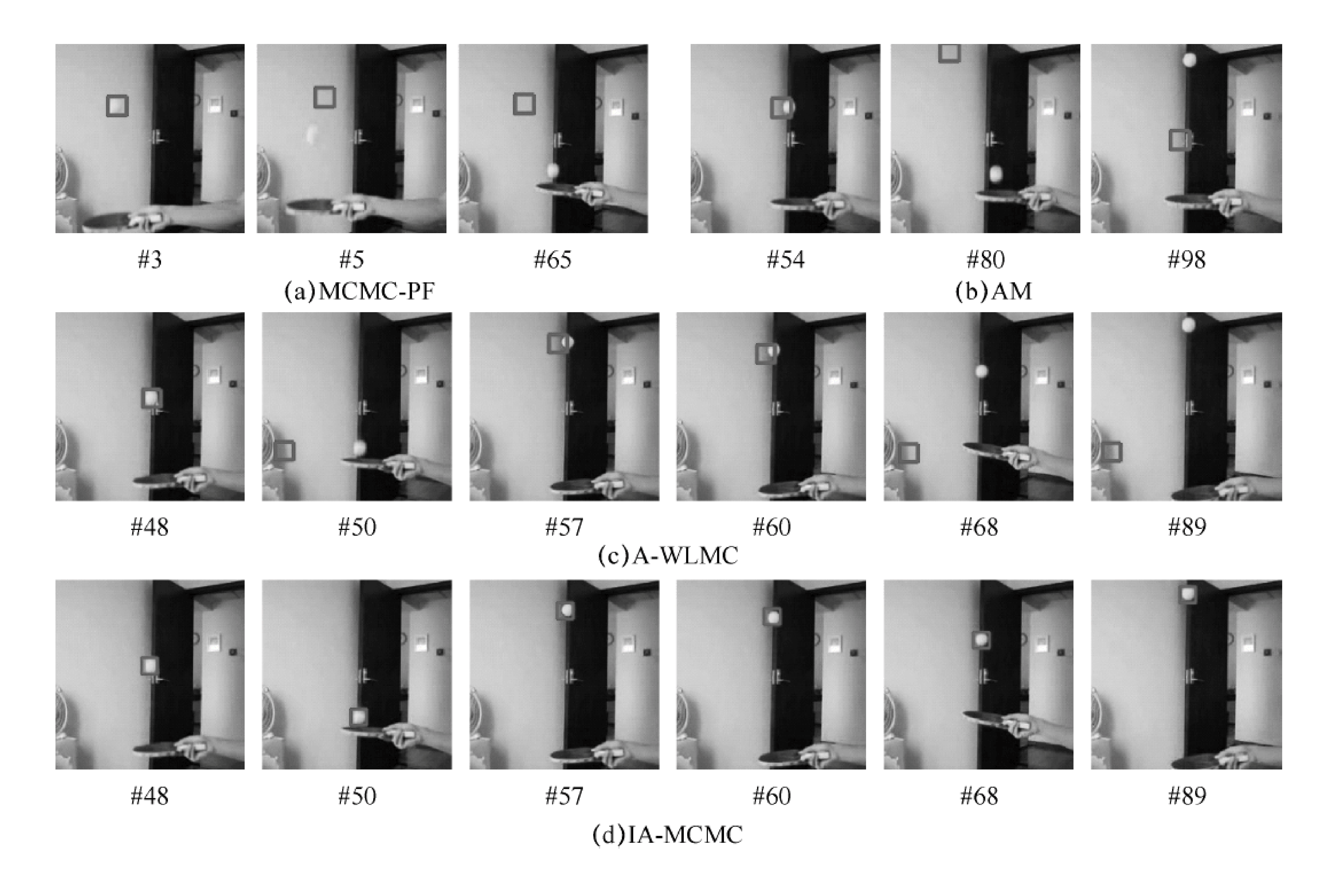


图 4-9　四种不同算法在 Pingpong 视频上的部分跟踪结果。MCMC-PF 和 AM 使用 1 000 个粒子，A-WLMC 和 IA-MCMC 使用 300 个粒子。

现跟踪漂移和失败。而仅使用了 300 个粒子，IA-MCMC 能够有效地跟踪到目标的极大不确定性。另一方面，同样使用 300 个粒子，A-WLMC 在一些帧中无法有效地估计目标的突变运动(例如＃26，＃33)。

5. 镜头切换＋低帧率视频

在这个实验中，我们定性地评估四种不同算法在一个编辑合成的视频 Walk 上的跟踪性能。在该合成视频的场景中，目标存在很严重的突变运动，它涉及两种类型的突变运动，即镜头切换和低帧率视频。该视频由三段不同视角拍摄的低帧率视频编辑合成的。此外，场景中多个行人来回走动，频繁地互相遮挡。部分跟踪结果如图 4-11 所示。从图 4-11 中可以看到，使用 300 个粒子，除了极少数帧(当目标几乎被其他目标完全遮挡时，例如＃100)，IA-MCMC 能在整个视频中实现有效地目标跟踪。同时我们可以看到，同样使用 300 个粒子，A-WLMC 在相当数量的视频帧中(当目标靠近或局部遮挡发生时)都出现跟踪失败的情况。另一方面，即便使用 600 个粒子，由于目标存在严重的运动不确定性，MCMC-PF 和 AM 算法频繁地出现了跟踪漂移和失败。



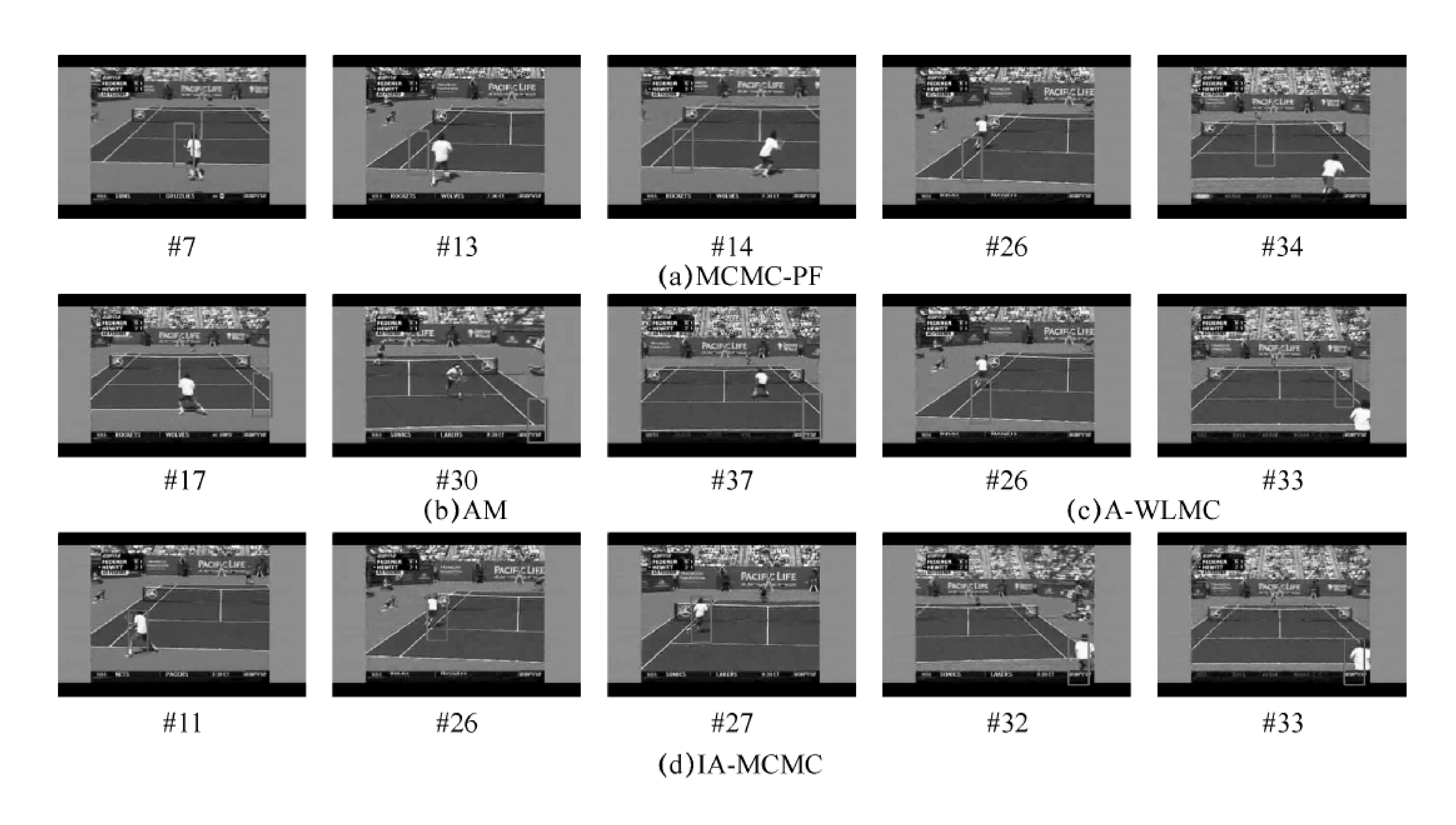


图 4-10　四种不同跟踪算法在 Tennis 视频上的部分跟踪结果。MCMC-PF 和 AM 使用 1 000 个粒子，A-WLMC 和 IA-MCMC 使用 300 个粒子。

4.4.3　定量比较

为量化地评估 IA-MCMC 跟踪算法的性能，我们在多个测试视频上与其他三种跟踪算法进行了定量的比较。我们首先定义跟踪失败，当估计的目标中心不在基准目标矩形(人为标注的)内时，就认为是跟踪失败。确切地说，在每一帧中只有当召回率(Recall)和准确率(Precision)都高于 50%时，才认为该帧的目标跟踪是成功的。

表 4-1 所示为四种不同跟踪算法在测试视频上的跟踪成功率。从表 4-1 中可以看到，即便使用 1 000 个粒子，MCMC-PF 和 AM 在这些存在突变运动的测试视频上表现出较差的跟踪性能，相对来说 AM 要稍好一些，这主要归功于其自适应的采样机制。然而由于缺乏能够避免采样陷入局部模式的有效机制，这两种算法在跟踪突变运动的目标时存在很大的困难。相比 MCMC-PF 和 AM 算法，A-WLMC 和 IA-MCMC 在使用较少粒子的情况下仍然在测试视频上获得了较高的跟踪成功率。此外，IA-MCMC 算法的跟踪成功率较之 A-WLMC 高出近 9 个百分点，而仅仅使用了其一半数量的粒子。

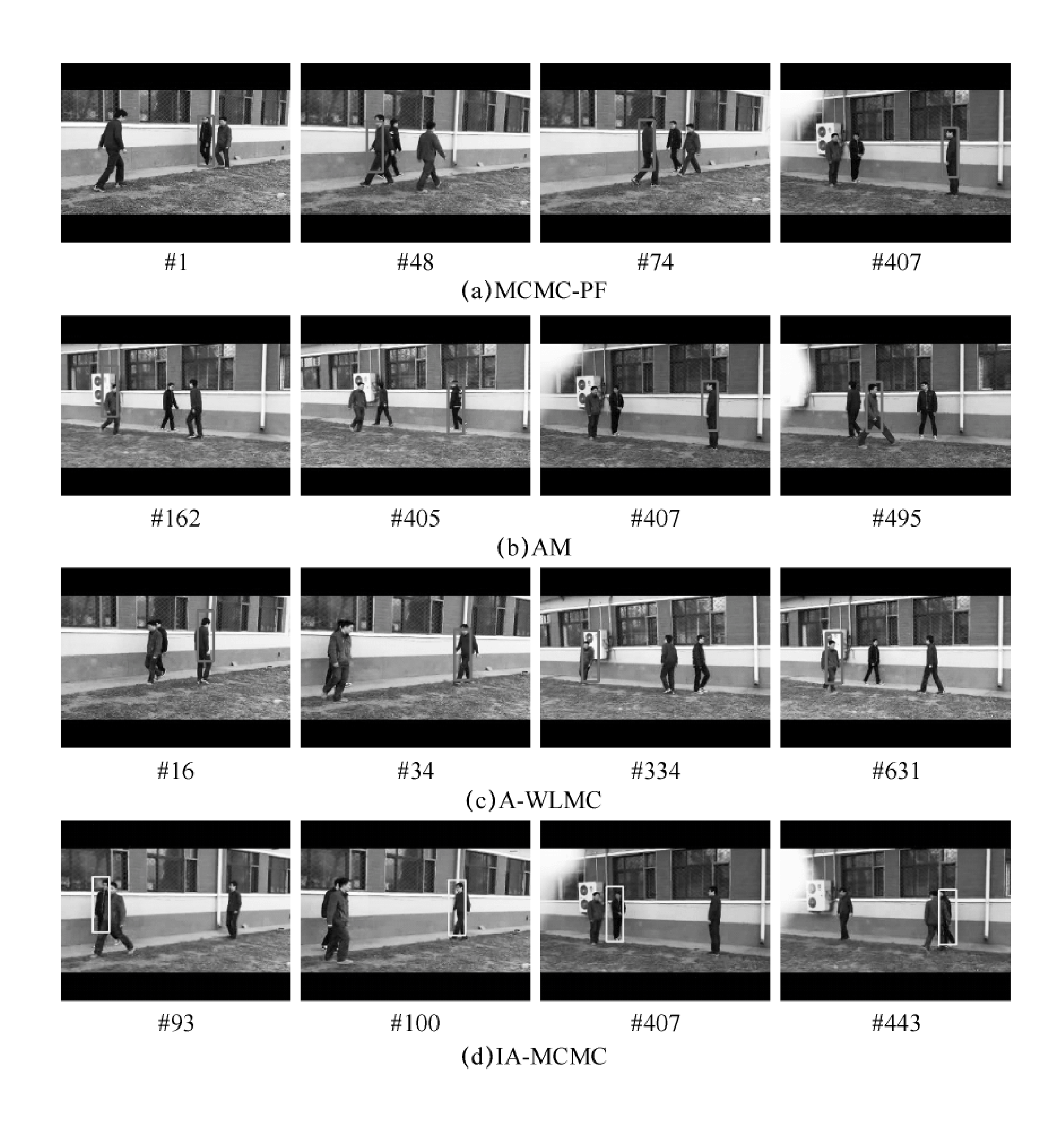


图 4-11　四种不同算法在 Walk 视频上的部分跟踪结果。MCMC-PF 和 AM 使用 600 个粒子，A-WLMC 和 IA-MCMC 使用 300 个粒子。

表 4-1　四种不同跟踪算法在测试视频上的跟踪成功率

方法	粒子数	跟踪成功率(%)	方法	粒子数	跟踪成功率(%)
MCMC-PF	1 000	68.3	A-WLMC	600	87.4
AM	1 000	74.1	IA-MCMC	300	96.1

此外，我们比较了四种算法在测试视频上的估计位置错误。图 4-12 所示为不同跟踪算法在两个测试视频上的逐帧位置错误比较，而相对位置错误(RPE)的统计见表 4-2。从表 4-2 中可以看到，无论是位置错误均值还是标准差，我们提出的 IA-MCMC 算法都明显小于其他三种算法。这表明了我们的算法在处理突变运动时能够获得更准确和稳定的跟踪性能。



最后，我们评估四种不同的跟踪算法在对低帧率视频跟踪时受帧率的影响情况。我们选择了一组低帧率视频：Tennis(10)，Tennis(15)，Tennis(20)，Tennis(30)，Tennis(35)。它们是由一段超过 700 帧的视频中按每 10～35 帧抽取 1 帧生成的低帧率视频序列。图 4-13 所示为四种不同跟踪算法在这个低帧率序列上跟踪成功率受帧率的影响。从图 4-13 中可以看到，即使使用少量的粒子，IA-MCMC 算法的跟踪性能明显优于其他三种算法。此外，IA-MCMC 和 A-WLMC 的跟踪性能受帧率的影响较小。而 MCMC-PF 和 AM 对帧率变化较为敏感，随着帧率的下降其跟踪性能也显著下降。

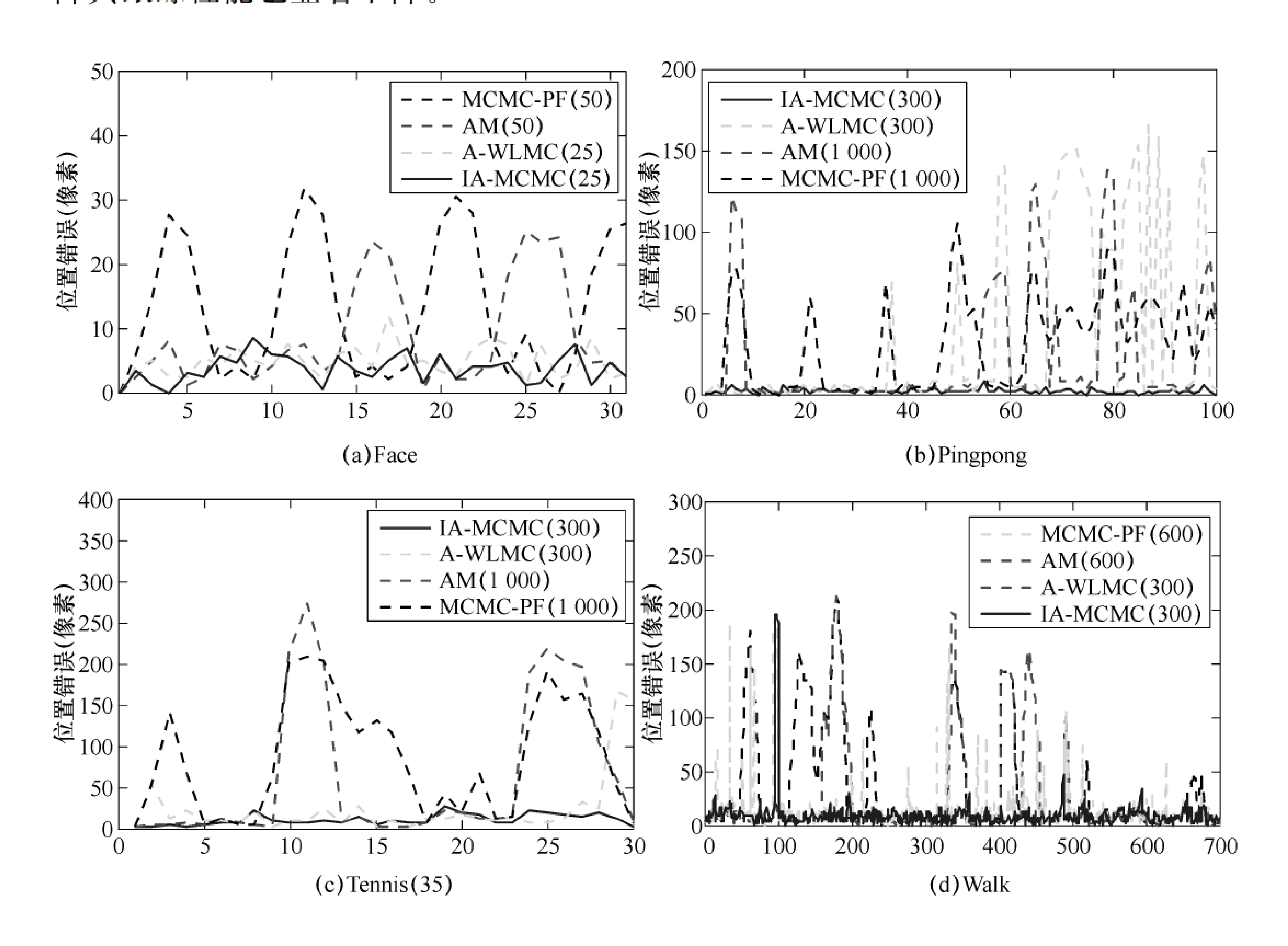


图 4-12　不同跟踪算法在四个测试视频上的跟踪精度比较。

4.4.4　讨论

以上实验验证了本章提出的 IA-MCMC 跟踪算法的有效性。当目标运动发生突变时，我们能够解释 MCMC-PF 和 AM 两种跟踪算法表现欠佳的主要原因，那就是它们缺乏避免采样陷入局部模式的有效机制即便使用大量的粒子，它们也易

于导致跟踪失败。另一方面，虽然 A－WLMC 算法具有处理突变运动的能力，但是它仍然需要相当数量的粒子实现在庞大的样本空间中进行充分的采样。不同于别的跟踪算法，IA-MCMC 利用马氏链上的样本信息进行深度自适应的 MCMC 采样，其中目标分布自适应用于避免采样陷入局部模式，而提议分布自适应则力图通过学习“最优的”提议分布来提高整体采样效率。IA-MCMC 可以看作一个嵌入数据挖掘模式的“采样—预测—采样”过程，它指导 MCMC 采样迅速地靠近或达到滤波分布的高置信度区域，从而实现有效地处理目标的突变运动。

表 4-2　不同跟踪算法在测试视频上的相对位置错误(RPE)

跟踪算法	Face		Pingpong		Tennis(35)		Walk	
	$\bar{d}$	d_σ	$\bar{d}$	d_σ	$\bar{d}$	d_σ	$\bar{d}$	d_σ
MCMC-PF	0.5997	0.4887	1.5571	1.3633	2.0175	1.7516	1.2710	2.0469
AM	0.3896	0.3589	1.1502	1.8803	1.3132	2.0012	0.9732	1.9215
A-WLMC	0.2234	0.1197	1.8041	2.7661	0.4867	0.8858	0.4575	1.0337
IA-MCMC	0.1706	0.1024	0.1396	0.0821	0.1716	0.1256	0.3316	0.7847

4.5　小　　结

本章提出了 IA-MCMC 采样算法实现运动突变的视觉跟踪任务。IA－MCMC 跟踪算法的关键环节是在 S-SAMC 跟踪框架下在线学习和估计“最优的”提议分布。为此，构建了一个基于密度网格的预测模型并将其嵌入到 MCMC 采样过程中，实现一个“采样—预测—采样”的过程。这种嵌入数据挖掘模式的 MCMC 采样算法能够显著地提高采样效率。IA-MCMC 算法能够有效地处理多种运动突变的视觉跟踪问题，包括目标的快速运动、目标动力学特征突变、低帧率视频以及综合多种运动突变的视频。全面的实验评估表明了与同类的主流跟踪算法相比它能获得更好的跟踪性能。

需要指出的是，本章主要关注运动突变的视觉跟踪问题，仅采用了一个简单的外观模型，并未考虑目标外观模型的自适应策略。即便如此，由于本章提出的跟踪算法是一个较为通用的跟踪方法，它易于集成更为有效的目标表示和可靠的外观模型（如判别外观模型[60, 62]）以进一步提高算法在复杂场景中的跟踪性能。

结　　论

视觉跟踪是计算机视觉研究的一个基本和关键的环节，长期以来它得到了广泛和深入的研究。近年来，随着相关学科领域理论和技术的发展，视觉跟踪研究在理论方法和实现技术方面得到了进一步的深入和发展，并且在智能交通、人机交互、视频监控和军事等领域得到了更为广泛的应用。在实际应用中，视觉跟踪方法通常需要处理一些复杂的情况，包括如何建模目标的运动变化，以有效地捕获目标运动的不确定性；如何设计有效的外观模型以适应外观的显著变化；如何设计稳健高效的视觉跟踪算法以满足实际的目标跟踪任务；等等。运动极大不确定性条件下的目标跟踪是视觉跟踪的一个难点，基于运动连续性假设的传统贝叶斯滤波跟踪方法在解决这一问题时具有很大的局限性。本书主要研究了自适应马氏链蒙特卡洛采样的理论方法和实现技术，以及在目标运动极大不确定性条件下的视觉跟踪方法。具体地：

(1) 研究了传统的 MCMC 采样算法在高维状态空间中采样时表现出混合性较差且易于陷入局部模式的基本原因；研究了当前主要的三类 A-MCMC 采样算法，即再生马氏链、提议分布自适应以及目标分布自适应；探讨了 MCMC 采样算法在计算机视觉中的应用，给出了一种结合局部优化与全局蒙特卡洛采样的目标轮廓近似表示算法。实验表明，与同类的主流多边形近似算法相比，本书研究的近似算法能取得更高的近似精度代价比。

(2) 研究了基于 SMC 采样和基于 MCMC 采样的视觉跟踪算法的局限性，将随机逼近蒙特卡洛(SAMC)采样方法引入到贝叶斯滤波跟踪框架，给出了一种序贯随机逼近蒙特卡洛(Sequential SAMC)跟踪算法，它是一个动态加权的重要性采样算法，能够有效避免粒子采样陷入局部模式。该算法是一个通用的视觉跟踪框架，易于集成不同的观测模型、目标表示以及提议分布，实现更为稳健的视觉跟踪。与同类的基于随机采样视觉跟踪框架(如 Bootstrap PF，MCMC-PF)相比，本书研

究的 Sequential SAMC 跟踪框架能够更为有效地处理复杂的贝叶斯滤波跟踪问题，同时还保持了同类跟踪框架所具有的可扩展性。

(3) 针对运动突变情况下的视觉跟踪问题，本书研究了一种强自适应的 MCMC(IA-MCMC)采样算法实现对复杂滤波分布的有效采样。它通过引入一个基于密度网格的在线数据挖掘模型到 SAMC 采样过程中，利用自适应马氏链上的历史样本信息，自适应地更新马尔科夫转移核。它通过两个相互协作的在线学习过程实现深度自适应的蒙特卡洛采样，其中一个学习过程通过在线估计态密度来逼近目标分布，另一个学习过程通过预测模型估计最佳的提议分布；该方法通过在能量空间的随机游走有效地避免采样陷入局部模式，同时能够通过自适应的采样机制加速收敛；实验表明，该算法在采样效率和跟踪性能方面明显地优于同类算法，并能够在多种复杂运动场景下(目标快速运动、目标动力学突变、镜头切换和低帧率视频)实现快速有效的目标跟踪。

复杂场景下的视觉跟踪是计算机视觉研究领域中一项困难的和充满挑战的研究方向。长期以来，国内外的学者对视觉跟踪进行了不懈地探索和研究，并取得了许多重要的理论成果和可靠实用的实践技术。即便如此，如何在复杂条件下实现稳健的视觉跟踪仍然是个极具挑战的研究课题，当前的视觉跟踪技术离实际应用需求还有相当大的距离。本书是在前人研究工作的基础上，对视觉跟踪领域中的一些关键问题进行了一些探索和研究，取得了一些研究成果。需要指出的，本书主要关注目标运动极大不确定性情况下的视觉跟踪问题，因而没有讨论目标外观的自适应。作者相信，当前成熟的外观模型(原生模型或判别模型)可以较为容易地集成到提出的跟踪框架中，以进一步提高跟踪算法的稳健性。

此外，自适应 MCMC 采样方法在最近几年取得了重要的理论进展，但该领域的理论研究和实现技术还有待进一步、深入地研究和探索。尽管如此，自适应 MCMC 采样方法已经在一些科学计算领域展示了其优越性。本书将其引入到视觉跟踪领域，做了一些有益的探索和研究，相信随着自适应 MCMC 采样理论研究的进一步深入与发展，将会进一步促进它在视觉跟踪研究乃至计算机视觉研究中的应用。

参 考 文 献

[1] 贾云得. 机器视觉[M]. 北京:科学出版社,2000.

[2] 侯志强,韩崇昭. 视觉跟踪技术综述. 自动化学报[J]. 2006,4(32),pp:603-617.

[3] W. Hu, T. Tan, L. Wang, S. Maybank. A survey on visual surveillance of object motion and behaviors. IEEE Trans. on SMC-C: Applications and Reviews [J]. 2004. 34(3), 334-352.

[4] M. Isard, A. Blake. CONDENSATION: Conditional density propagation for visual tracking. IJCV [J]. 1998, 29(1): 5-28.

[5] Z. Khan, T. Balch, F. Dellaert. An mcmc-based particle filter for tracking multiple interacting targets. ECCV [C], 2004, 279-90.

[6] A. Yilmaz, O. Javed, M. Shah. Object tracking: A survey. ACM Computing Surveys [J]. 2006, 38(4).

[7] C. J. Veenman, M. J. T. Reinders, E. Backer. Resolving motion correspondence for densely mo- ving points. IEEE Trans. on PAMI [J]. 2001, 23(1), 54-72.

[8] D. Serby, E. K. Meier, L. V. Gool. Probabilistic object tracking using multiple features. ICPR [C], 2004, 184-187.

[9] D. Comaniciu, V. Ramesh, P. Meer. Kernel-based object tracking. IEEE Trans. on PAMI [J]. 2003, 25(5), 564-577.

[10] A. Yilmaz, X. Li, M. Shah. Contour-based object tracking with occlusion handling in video acquired using mobile cameras. IEEE Trans. on PAMI [J]. 2004, 26(11), 1531-1536.

[11] T. F. Cootes, G. J. Edwards, C. J. Taylor. Active appearance models. IEEE Trans. on PAMI [J]. 2001, 23(6), 681-685.

[12] J. Canny. A computational approach to edge detection. IEEE Trans. on PAMI [J]. 1986, 8(6), 679-698.

[13] G. Hua, Y. Wu. Multi-scale visual tracking by sequential belief propagation. CVPR [C], 2004, 826-833.

[14] B. D. Lucas, T. Kanade. An iterative image registration technique with an application to stereo vision. IJCAI [C], 1981, 674-679.

[15] B. K. P. Horn, B. G. Schunck. Determining optical flow. Artificial Intelligence [J]. 1981, 17(1), 185-203.

[16] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection. CVPR [C], 2005, 886-893.

[17] T. Ahonen, A. Hadid, M. Pietikinen. Face recognition with local binary patterns. ECCV [C], 2004, 469-481.

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. IJCV [J]. 2004, vol. 60, 91 – 110.

[19] K. Tieu, P. Viola. Boosting image retrieval. IJCV [J]. 2004, 56(1), 17-36.

[20] J. Wang, X. Chen, W. Gao. Online selecting discriminative tracking features using particle filter. CVPR [C], 2005, vol. 2, 1037-1042.

[21] P. Fieguth, D. Terzopoulos. Color-based tracking of heads and other mobile objects at video frame rates. CVPR [C], 1997, 21-27.

[22] E. Maggio, A. Cavallaro. Accurate appearance-based Bayesian tracking for maneuvering targets. CVIU [J]. 2009, 113(4), 544-555.

[23] S. C. Zhu, A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. IEEE Trans. on PAMI [J]. 1996, 18(9), 884-900.

[24] S. J. McKenna, Y. Raja, S. Gong. Tracking colour objects using adaptive mixture models. IVC [J]. 1999, 17(3), 225-231.

[25] A. Elgammal, R. Duraiswami, D. Harwood, L. S. Davis. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proceedings of the IEEE [J]. 2002, 90(7), 1151-1163.

[26] M. J. Black, A. D. Jepson. EigenTracking: Robust matching and tracking of articulated objects using a view-based representation. IJCV [J]. 1998, 26(1), 63-84.

[27] S. Avidan. Support vector tracking. IEEE Trans. on PAMI [J]. 2004, 26(8), pp: 1064-1072.



[28] J. Shi, C. Tomasi. Good features to track. CVPR [C], 1994, pp: 593-600.

[29] S. Baker, I. Matthews. Lucas-Kanade 20 years: A unifying framework. IJCV [J]. 2004, 56(3), 221-255.

[30] I. Matthews, T. Ishikawa, S. Baker. The template update problem. IEEE Trans. on PAMI [J]. 2004, 26(6), pp: 810-815.

[31] G. D. Hager, P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. on PAMI [J]. 1998, 20 (10), 1025-1039.

[32] H. T. Nguyen, A. W. M. Smeulders. Fast occluded object tracking by a robust appearance filter. IEEE Trans. on PAMI [J]. 2004, 26 (8), 1099-1104.

[33] A. D. Jepson, D. J. Fleet, T. F. El-Maraghi. Robust online appearance models for visual tracking. IEEE Trans. on PAMI [J]. 2003, 25 (10), 1296-1311.

[34] B. Han, L. Davis. On-line density-based appearance modeling for object tracking. ICCV [C], 2005, 1492-1499.

[35] S. Birchfield. Elliptical head tracking using intensity gradients and color histograms. CVPR [C], 1998, 232-237.

[36] D. Comaniciu, V. Ramesh, P. Meer. Real-time tracking of non-rigid objects using Mean Shift. CVPR [C], 2000, 142-149.

[37] R. T. Collins. Mean-shift blob tracking through scale space. CVPR [C], 2003, 234- 240.

[38] P. Pérez, C. Hue, J. Vermaak, M. Gangnet. Color-nased probabilistic tracking. ECCV [C], 2002, 661-675.

[39] K. Nummiaro, E. Koller-Meier, L. V. Gool. An adaptive color-based particle filter. IVC [J]. 2003, 21(1), 99-110.

[40] F. Porikli. Integral histogram: A fast way to extract histograms in Cartesian spaces. CVPR [C], 2005, 829-836.

[41] A. Adam, E. Rivlin, I. Shimshoni. Robust fragments-based tracking using the integral histogram. CVPR [C], 2006, 798-805.

[42] Z. Zivkovic, B. Krose. An EM-like algorithm for color-histogram-based object tracking. CVPR [C], 2004, 798-803.

[43] G. D. Hager, M. Dewan, C. V. Stewart. Multiple kernel tracking with SSD. CVPR [C], 2004, 790-797.

[44] Z. Fan, Y. Wu, M. Yang. Multiple collaborative kernel tracking. CVPR [C], 2005, 502-509.

[45] M. Dewan, G. D. Hager. Toward optimal kernel-based tracking. CVPR [C], 2006, 618-625.

[46] Z. Fan, M. Yang, Y. Wu, G. Hua, T. Yu. Efficient optimal kernel placement for reliable visual tracking. CVPR [C], 2006, 658-665.

[47] D. Ross, J. Lim, M. -H. Yang. Adaptive probabilistic visual tracking with incremental subspace update. ECCV [C], 2004, 470-482.

[48] X. Li, W. Hu, Z. Zhang, X. Zhang, M. Zhu, J. Cheng. Visual tracking via incremental Log- Euclidean Riemannian subspace learning. CVPR [C], 2008.

[49] Y. Wu, J. Cheng, J. Wang, H. Lu. Real-time visual tracking via incremental covariance tensor learning. ICCV [C], 2009, 1631-1638.

[50] X. Li, W. Hu, Z. Zhang, X. Zhang. Robust visual tracking based on an effective appearance model. ECCV [C], 2008, 396-408.

[51] R. T. Collins, Y. Liu. On-line selection of discriminative tracking features. ICCV [C], 2003, 346-352.

[52] Z. Yin, R. Collins. Spatial divide and conquer with motion cues for tracking through clutter. CVPR [C], 2006, 570-577.

[53] J. Wang, X. Chen, W. Gao. Online selecting discriminative tracking features using particle filter. CVPR [C], 2005, 1037-1042.

[54] M. Yang, Z. Fan, J. Fan, Y. Wu. Tracking nonstationary visual appearances by data-driven ada- ptation. IEEE Trans. on Image Processing [J]. 2009, 18(7), 1633-1644.

[55] S. Avidan. Ensemble tracking. CVPR [C], 2005, 494-501.

[56] H. Grabner, H. Bischof. On-line boosting and vision. CVPR [C], 2006, 260-267.

[57] F. Tang, S. Brennan, Q. Zhao and H. Tao. Co-tracking using semi-supervised support vector machines. ICCV [C], 2007, 1-8.



[58] R. Liu, J. Cheng, H. Lu. A robust boosting tracker with minimum error bound in a co-training framework. ICCV [C], 2009, 1459-15568.

[59] T. Woodley, B. Stenger and R. Cipolla. Tracking using online feature selection and a local genera-tive model. BMVC [C], 2007.

[60] H. Grabner, C. Leistner and H. Bischof. Semi-supervised on-line boosting for robust tracking. ECCV [C], 2008, 234-247.

[61] P. K. Mallapragada, R. Jin, A. K. Jain, Y. Liu. SemiBoost: Boosting for semi-supervised learning. IEEE Trans. on PAMI [J], 2009, 31(11), 2000-2014.

[62] B. Babenko, M. -H. Yang and S. Belongie. Visual tracking with online multiple instance leaning. CVPR [C], 2009, 983-990.

[63] P. Viola, J. C. Platt, C. Zhang. Multiple instance boosting for object detection. NIPS [C], 2005, 1417-1426.

[64] B. Zeisl, C. Leistner, A. Saffari, H. Bischof. On-line semi-supervised multiple-instance boos- ting. CVPR[C], 2010, 1879-1886.

[65] Y. Cheng. Mean Shift, mode seeking, clustering. IEEE Trans. on PAMI [J]. 1995, 17(8), 790- 799.

[66] D. Comaniciu, P. Meer. Mean Shift: A robust approach toward feature space analysis. IEEE Trans. on PAMI [J]. 2002, 24(5), 603-619.

[67] M. Fashing, C. Tomasi. Mean shift is a bound optimization. IEEE Trans. on PAMI [J]. 2005, 27(3), 471-474.

[68] C. Shan, T. Tan, Y. Wei. Real-time hand tracking using a Mean Shift embedded particle filter. Pattern Recognition [J], 2007, 40(1), 1958-1970.

[69] C. Shen, M. J. Brooks, A. Hengel. Fast global kernel density mode seeking with application to localization and tracking. ICCV [C], 2005, 1516-1523.

[70] R. E. Kalman. A new approach to linear filtering and prediction problems. Journal of Basic Engi- neering [J]. 1960.

[71] M. S. Arulampalam, S. Maskell, N. Gordon, T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. on Signal Processing. 2002, 50(2), 174- 188.

[72] A. Doucet, S. Godsill and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and Computing [J]. 2000, vol. 10, 197-208.

[73] M. Isard, A. Blake. ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework. ECCV [C], 1998, 893-908.

[74] J. S. Liu, R. Chen, T. Logvinenko. A theoretical framework for sequen tial importance sampling and resampling. Sequential Monte Carlo Methods in Practice [M], New York: Springer-Verlag, 2001.

[75] W. R. Gilks, S. Richardson, D. J. Spiegelhalter. Markov Chain Monte Carlo in practice [M]. New York: Chapman and Hall, 1996.

[76] B. Han, Y. Zhu, D. Comaniciu, L. Davis. Kernel-based Bayesian filtering for object tracking. CVPR [C], 2005, 227-234.

[77] G. Zhu, D. Liang, Y. Liu, Q. Huang, W. Gao. Improving particle filter with support vector regression for efficient visual tracking. ICIP [C]. 2005, 422-425.

[78] Y. Wu, J. Fan. Contextual flow. CVPR [C]. 2009, 33-40.

[79] T. J. Cham, J. M. Rehg. A Multiple hypothesis approach to figure tracking. CVPR [C], 1999, 239-245.

[80] J. Vermaak, A. Doucet, P. Perez. Maintaining multi-modality through mixture tracking. ICCV [C], 2003, 1110-1116.

[81] A. Banerjee, P. Burlina. Efficient particle filtering via sparse kernel density estimation. IEEE Trans. on Image Processing [J]. 2010, 19 (9), 2480-2490.

[82] J. Weston. Support vector density estimation. Advances in Kernel Methods [M], MIT Press, 1998.

[83] J. Deutcher, A. Blake, I. Reid. Articulated body motion capture by annealed particle filtering. CVPR [C], 2000, 2126-2133.

[84] K. Choo, D. Fleet. People tracking using hybrid Monte Carlo filtering. ICCV [C], 2001, 321-328.

[85] T. Zhao, R. Nevatia. Tracking multiple humans in crowded environment. CVPR [C], 2004, 406-413.



[86] M. Isard, A. Blake. A Mixed-state condensation tracker with automatic model switching. ICCV [C], 1998, 107-112.

[87] B. North, A. Blake. Learning dynamical models by expectation maximization. ICCV [C], 1998, 384-389.

[88] J. Sullivan, A. Blake, M. Isard, J. MacCormick. Object localization by Bayesian correlation. ICCV [C], 1999, 1068-1075.

[89] Y. Li, H. Ai, T. Yamashita, S. Lao, M. Kawade. Tracking in low frame rate video: A cascade particle filter with discriminative observers of different life-spans. IEEE Trans. on PAMI [J], 2008. 30(10), 1728-1740.

[90] J. Kwon, K. Lee. Tracking of abrupt motion using Wang-Landau Monte Carlo estimation. ECCV [C], 2008, 387-400.

[91] F. Wang, D. P. Landau. Efficient multiple-range random-walk algorithm to calculate the density of states. Physical Review Letters [J]. 2001, 86 (10), 2050-2053.

[92] B. A. Berg, T. Neuhaus. Multicanonical algorithms for 1st-order phase-transitions. Physics Letters B [J]. 1991, vol. 267, 249 – 253.

[93] F. Liang, C. Liu, R. J. Carroll. Stochastic approximation in Monte Carlo computation. J. Amer. Statist. Assoc. [J]. 2007, 102(477), 305 – 320.

[94] X. Zhou, Y. Lu. Abrupt motion tracking via adaptive stochastic approximation Monte Carlo sampling. CVPR [C], 2010, 1847-1854.

[95] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. Biometrika [J]. 1970, 57 (1), 97-109.

[96] S. Geman, D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. on PAMI [J]. 1984, 6(6), 721-741.

[97] J. S. Liu, F. Liang, W. H. Wong. A theory for dynamic weighting in Monte Carlo. J. Amer. Statist. Assoc. [J]. 2001, vol. 96, 561-573.

[98] F. Liang. Dynamically weighted importance sampling in Monte Carlo computation. J. Amer. Statist. Assoc. [J]. 2002, vol. 97, 807-821.

[99] W. Gilks, G. O. Roberts, S. Sahu. Adaptive Markov chain Monte Carlo through regeneration. J. Amer. Statist. Assoc. [J]. 1998, vol. 93, 1045-1054.

[100] Y. Atchadé, J. S. Rosenthal. On adaptive Markov chain Monte Carlo algorithms. Bernoulli [J]. 2005, vol. 11, 815-828.

[101] H. Haario, E. Saksman, J. Tamminen. An adaptive metropolis algorithm. Bernoulli [J]. 2001, 7(2), 223-242.

[102] G. O. Roberts, J. S. Rosenthal. Coupling and ergodicity of adaptive Markov chain Monte Carlo algorithms. J. Appl. Probab. [J]. 2007, 44 (2), 458-475.

[103] G. O. Roberts, J. S. Rosenthal. Examples of adaptive MCMC. J. Comput. Graph. Statist. [J]. 2009, 18(2), 349-367.

[104] A. E. Gelfand, S. K. Sahu. On Markov chain Monte Carlo acceleration. Journal of Computa- tional and Graphical Statistics [J]. 1994, vol. 3, 261-276.

[105] W. J. Browne, D. Draper. Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. Computational Statistics [J]. 2000, vol. 15, 391-420.

[106] A. E. Brockwell, J. B. Kadane. Identification of regeneration times in MCMC simulation, with application to adaptive schemes. J. Comput. Graph. Stat. [J], 2005, vol. 14, 436 - 458.

[107] C. Andrieu, C. P. Robert. Controlled MCMC for optimal sampling. MCMC Preprint Service. 2002, http://www.ceremade.dauphine.fr/~xian/control.ps.gz.

[108] G. Casella, C. P. Robert. Rao-Blackwellisation of sampling schemes. Biometrika [J]. 1996, 83(1), 81-94.

[109] R. Y. Rubinstein. Simulation and the Monte Carlo method [M]. New York: John Wiley and Sons. 1981.

[110] J. Geweke. Bayesian inference in econometric models using Monte Carlo integration. Econome- trica [J]. 1989, vol. 24, pp: 1317-1399.

[111] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller. Equation of state calculaiton by fast computing machines. Journal of Chemical Physics [J]. 1953, vol. 21, 1087- 1091.

[112] J. S. Liu, F. Liang, W. H. Wong. The multiple-try method and local optimization in Metropolis sampling. J. Amer. Statist. Assoc. [J]. 2000, vol. 95, 121-134.



[113] C. P. Robert, G. Casella. Monte Carlo statistical methods (2nd edition) [M]. New York: Springer. 2004.

[114] C. Ritter, M. A. Tanner. Facilitating the Gibbs sampler: the Gibbs stopper and the griddy-Gibbs sampler. J. Amer. Statist. Assoc. [J]. 1992, 87 (419), 861-868.

[115] H. Robbins, S. Monro. A stochastic approximation method. The Annals of Mathematical Statistics [J]. 1951, vol. 22, 400-407.

[116] C. Yang. Recurrent, ergodic properties of adaptive MCMC. Preprint, 2008.

[117] L. Holden, R. Hauge, M. Holden. Adaptive independent Metropolis-Hastings. The Annals of Applied Probability [J]. 2009, 19(1), 395-413.

[118] Y. Ren, Y. Ding, F. Liang. Adaptive evolutionary Monte Carlo algorithm for optimization with applications to sensor placement problems. Statistics and Computing [J]. 2008, 18(4), 375- 390.

[119] A. Gelman, G. Roberts, W. Gilks. Efficient Metropolis jumping rules. In: Bayesian Statistics [M]. vol. 5. New York: Oxford University Press, 1995.

[120] C. Andrieu, É. Moulines. On the ergodicity properties of some adaptive MCMC algorithms. Ann. Appl. Probab. [J]. 2006, 16(3), 1462-1505.

[121] G. O. Roberts, J. S. Rosenthal. Optimal scaling for various Metropolis-Hastings algorithms. Stat. Sci. [J], 2001, vol. 16, 351-367.

[122] F. Liang. Generalized Wang-Landau algorithm for Monte Carlo computation. J. Amer. Statist. Assoc. [J]. 2005, vol. 100, 1311-1327.

[123] Q. Yan, J. J. de Pablo. Fast Calculation of the density of states of a fluid by Monte Carlo simulations. Physics Review Letters [J]. 2003, 90, 035701.

[124] C. Andrieu, É. Moulines, P. Priouret. Stability of stochastic approximation under verifiable conditions. SIAM Journal of Control and Optimization [J]. 2005, vol. 44, 283 - 312.

[125] X, Zhou, Y. Lu. Efficient polygonal approximation of digital curves via Monte Carlo optimization. ICPR [C], 2010, 3513-3516.

[126] X. Zhou, Y. Lu. Polygonal approximation of digital curves using adaptive MCMC sampling. ICIP [C], 2010, 2753-2756.



[127] M. K. Cowles, B. P. carlin. Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Amer. Statist. Assoc. [J]. 1996, 91 (434), 883-904.

[128] S. B. Brooks, G. O. Roberts. Assessing convergence of Markov chain Monte Carlo algorithms. Statistics and Computing [J]. 1999, vol. 8, 319-335.

[129] O. Cappé, A. Guillin, J-M. Marin, C. P. Robert. Population Monte Carlo. J. Comput. Graph. Statist. [J]. 2004, 13(4), 907-929.

[130] A. Gelman, D. B. Rubin. Inference from iterative simulation using multiple sequences. Statis- tical Science [J]. 1992, vol. 7, 457-511.

[131] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin. Bayesian Data Analysis (2nd edition) [M]. London: Chapman & Hall, 2004.

[132] B. K. Ray, K. S. Ray. A new split-and-merge technique for polygonal approximation of chain coded curves. Pattern Recognition Letters [J]. 1995, vol. 16, 161-169.

[133] A. Masood. Dominant point deletion by reverse polygonization of digital curves. Image and Vision Computing [J]. 2008, vol. 26, 702-715.

[134] J. C. Perez, E. Vidal. Optimum polygonal approximation of digitized curves. Pattern Recog- nition Letters [J]. 1994, vol. 15, 743-750.

[135] S. Y. Ho, Y. C. Chen. An efficient evolutionary algorithm for accurate polygonal approximation. Pattern Recognition [J]. 2001, vol. 34, 2305-2317.

[136] B. Wang, H. Shu, L. Luo. A genetic algorithm with chromosome-repairing for min-# and min-ϵ polygonal approximation of digital curves. J. Vis. Commun. Image R. [J]. 2009, vol. 20, 45-56.

[137] B. Wang, H. Shu, C. Shi, L. Luo. A novel stochastic search method for polygonal approximation problem. Neurocomputing [J]. 2008, vol. 71, 3216-3223.

[138] P. Y. Yin. Ant colony search algorithms for optimal polygonal approximation of plane curves. Pattern Recognition [J]. 2003, vol. 36, 1783- 1997.

[139] P. Y. Yin. A discrete particle swarm algorithm for optimal polygonal approximation of digital curves. J. Vis. Commun. Image R. [J]. 2004, vol. 15, 241-260.



[140] D. Sarkar. A simple algorithm for detection of significant vertices for polygonal approximation of chain-coded curves. Pattern Recogniton Letters [J]. 1993, vol. 14, 959-964.

[141] D. J. Wales, J. P. K . Doye. Global optimization by basin-hopping and the lowest energy structure of Lennard-Jones clusters containing up to 110 atoms. J. Phys. Chem. [J]. 1997, vol. 101, 5111-5116.

[142] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi. Optimization by simulated annealing. Science [J]. 1983, 220(4598), 671-680.

[143] R. van der Merwe, A. Doucet, J. F. G. de Freitas, E. Wan. The unscented particle filter. Technical Report CUED/F-INFENG/TR 380, Cambridge University Engineering Department, Cambridge, England, 2000.

[144] O. Cappe, S. J. Godsill, E. Moulines. An overview of existing methods and recent advances in sequential Monte Carlo. Proceedings of the IEEE [J]. 2007, 95(5), 899-924.

[145] F. Liang. On the use of stochastic approximation Monte Carlo for Monte Carlo integration. Statistics and Probability Letters [J]. 2009, vol. 79, 581-587.

[146] Y. Chen, L. Tu. Density-based clustering for real-time stream data. ACM SIGKDD [C]. 2007, 133-142.

[147] http://vision.stanford.edu/~birch/headtracker/seq/.